Biochemical Characterization of FAN1 Nuclease

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von

Julia Pizzolato

von

Zürich, ZH

Promotionskomitee

Prof. Dr. Josef Jiricny (Vorsitz, Leitung der Dissertation)
Prof. Dr. Alessandro A. Sartori
Prof. Dr. Orlando D. Schärer
Prof. Dr. Lumir Krejci

Zürich, 2016

Table of contents

Table of contents **1**

Zusammenfassung **3**

Summary **4**

1 Introduction **5**

1.1 DNA damage and repair 5

1.2 Replication and replication stress 6

1.2.1 Eukaryotic DNA replication 6
1.2.2 Replication stress 8
1.2.3 Recovery of stalled replication forks 9

1.3 Interstrand cross-link repair (ICL repair) 10

1.3.1 Overview of the ICL repair pathway 10
1.3.2 Types of ICL inducing agents 11
1.3.3 *Fanconi anemia* proteins and pathway 12
1.3.4 Unhooking step in ICL repair 15
1.3.5 Translesion polymerases and their roles in ICL repair 18
1.3.6 Homologous recombination and ICL repair 20
1.3.7 Nucleotide excision repair and ICL repair 24
1.3.8 ICL repair in the *Xenopus* system 25
1.3.9 Replication traverse of ICLs 26
1.3.10 Replication-independent ICL repair 27
1.3.11 ICL repair by glycosylases 28
1.3.12 ICL repair in yeast 29

1.4 FEN1 30

1.5 FAN1 31

1.5.1 FAN1 discovery 31
1.5.2 Roles of FAN1 outside of ICL repair 32
1.5.3 FAN1 in *Schizosaccharomyces pombe* 32
1.5.4 FAN1 structure 33
1.5.5 FAN1 in diseases 35

1.6 My aim 35

2 Results **37**

2.1 FANCD2-associated Nuclease 1, but not Exonuclease 1 or Flap Endonuclease 1, is able to unhook DNA interstrand cross-links *in vitro* 37

2.2 Additional Observations 48

2.2.1 5' Flap cleavage activity in human TSCRE2 extracts 48

2.2.2 FEN1 is responsible for almost all 5' flap cleavage in chicken cell extracts, but FAN1 activity becomes detectable upon overexpression 49

2.2.3 FEN1 depletion sensitizes chicken cells to MMC and this effect can be rescued by FAN1 overexpression 51

2.2.4 FAN1 protein levels in human cells are higher in the absence of FEN1 protein 52

3 Discussion 54

3.1 *In vitro* unhooking activity 54

3.2 *In vivo* implications of FAN1 unhooking activity 55

3.3 Evidence for FAN1 involvement in processes other than ICL unhooking 56

3.3.1 Replication fork recovery 57

3.3.2 Homologous recombination 57

3.4 FAN1 and FEN1 58

3.5 FAN1 and SNM1A 59

3.6 Potential parallel pathways to repair ICLs 59

4 Conclusions and perspectives 61

5 Materials and Methods to section 2.2 62

5.1 Substrates for endonuclease assays 62

5.2 Culturing of DT40 chicken cells 62

5.3 Culturing of TSCER2 cell lines and nuclear extract preparation 62

5.4 FAN1 overexpression in DT40 FEN1 KO cells 62

5.5 Preparation of whole cell extracts from DT40 cells 62

5.6 Nuclease assays with DT40 extracts 63

5.7 Methylcellulose survival assay with DT40 cells 63

5.8 quantitative real-time (RT) PCR 63

5.9 Antibodies 63

5.10 siRNAs 64

6 References 65

7 Acknowledgements 76

8 Curriculum Vitae 77

Zusammenfassung

DNS-Reparaturmechanismen sind essentiell, um die genomische Stabilität aufrecht zu erhalten und dadurch auch um Krebs entgegenzuwirken. DNS Interstrangvernetzungen (ICLs), die beide DNS-Stränge kovalent miteinander verbinden, gehören zu den gefährlichsten DNS Schäden, da sie sowohl die Transkription als auch die Replikation blockieren. In höheren Eukaryoten werden ICLs auf eine komplizierte Art und Weise repariert und viele verschiedene DNS-Reparaturmechanismen sind daran beteiligt, wie Transläsionssynthese (TLS), Nukleotidexzisionsreparatur (NER) und homologe Rekombination (HR). Proteine der Fanconi-Anämie Komplementationsgruppe werden für die Koordination der Reparatur benötigt. Patienten mit Mutationen in Genen, welche für diese Proteine kodieren leiden an der schwerwiegenden, genetischen Krankheit Fanconi-Anämie. Der ICL-Reparaturweg beginnt mit dem Aushaken der Vernetzung nachdem eine Replikationsgabel davor stecken geblieben ist, wobei Nukleasen auf beiden Seiten davon einschneiden um sie zu entfernen.

FAN1 (FANCD2-associated nuclease 1) wurde in vier unabhängigen Forschungsgruppen als ein wichtiges Protein für die Reparatur von ICLs identifiziert. Es verleiht Zellen Resistenz gegen ICL-induzierende Chemikalien, wie zum Beispiel MMC (Mitomycin C) und Cisplatin. Mit Hilfe von *in vitro* Versuchen wurde gezeigt, dass FAN1 sowohl eine Endo- als auch eine Exonuklease-Aktivität besitzt. Als Endonuklease schneidet es präferentiell sogenannte 5' Flap-Strukturen. Die Exonuklease-Aktivität von FAN1 zeigt eine 5'->3' Orientierung auf.

Mein Ziel war es, mehr über die Funktion von FAN1 im Zusammenhang mit der Reparatur von ICLs zu erfahren. Ich machte von synthetischer DNS mit ICLs Gebrauch, um zu sehen ob und wie FAN1 in der Lage ist DNS um die ICL herum abzubauen. Dadurch konnte ich zeigen, dass FAN1 tatsächlich die Vernetzung aushaken kann, indem es auf beiden Seiten, also 5' und 3' davon schneidet. Dies war nicht der Fall, als ich zwei andere Nukleasen mit ähnlichen Spezifitäten, EXO1 und FEN1, benutzte. Auch mit lebenden Zellen konnte ich zeigen, dass FAN1 involviert ist in der Induzierung von Doppelstrangbrüchen in Abhängigkeit von ICLs, was auch für eine Rolle im Aushakungsschritt spricht.

FAN1 ist ein potentielles Tumor-Suppressor-Gen und ICL-induzierende Chemikalien werden oft in der Chemotherapie verwendet. Deshalb könnte mehr Wissen auf dem Gebiet potentiell die Krebstherapie vereinfachen.

Summary

DNA repair mechanisms are essential to maintain genomic stability and therefore to prevent cancer. DNA interstrand cross-links (ICLs), which covalently link both strands of the DNA double helix, are among the most cytotoxic DNA lesions due to their potential to block both transcription and replication. In higher eukaryotes, ICLs are repaired in a complex manner, involving several different DNA repair mechanisms, such as translesion synthesis, nucleotide excision repair and homologous recombination. The repair is coordinated by proteins of the FA complementation group, which, when mutated, contribute to the severe genetic disorder *Fanconi anemia*. After replication fork stalling at the site of the cross-link, the first step of ICL repair consists of lesion unhooking, where nucleases cleave at both sides of the cross-link in order to remove it.

FAN1 (FANCD2-associated nuclease 1), which was independently identified in four different studies, possesses a conserved role in ICL repair, conferring resistance towards ICL inducing agents such as MMC (mitomycin C) or cisplatin. In *in vitro* assays, FAN1 was shown to contain both endo- and exonuclease activities, the former exhibiting a preference for 5' flap DNA structures and the latter a 5'->3' polarity.

My aim was to elucidate the function of FAN1 in the context of ICL repair. Specifically, I made use of synthetic DNA structures containing ICLs to assess the ability of FAN1 to unhook and degrade DNA across an ICL. I could show that FAN1 indeed is able to unhook ICLs *in vitro* by making incisions both 5' and 3' of them. This was not the case when I used two other nucleases with similar substrate specificities, EXO1 and FEN1. Using living cells, I could further show that the induction of ICL-dependent double-strand breaks is dependent on FAN1, which is consistent with a role in unhooking.

FAN1 is a potential tumor suppressor gene and ICL inducing agents are often used in cancer chemotherapy. Therefore, better understanding of ICL processing could eventually lead to improved treatments.

1 Introduction

1.1 DNA damage and repair

Our DNA is under constant attack of endogenous and exogenous agents that damage it in various ways (1). If the DNA is left unrepaired, this leads to mutations and eventually to the development of cancer. It can happen that genes important for the survival, proliferation capacity or growth of the cells are hyperactivated or that genes important to counteract these functions are by chance disrupted or downregulated. Furthermore, a gene modified by a point mutation can lead to gain of function of the gene product that can be beneficial for the survival of the cell and transform it in this way. These and other mechanisms ultimately lead to the formation of cancer, where cells start to proliferate in an uncontrolled fashion (2,3).

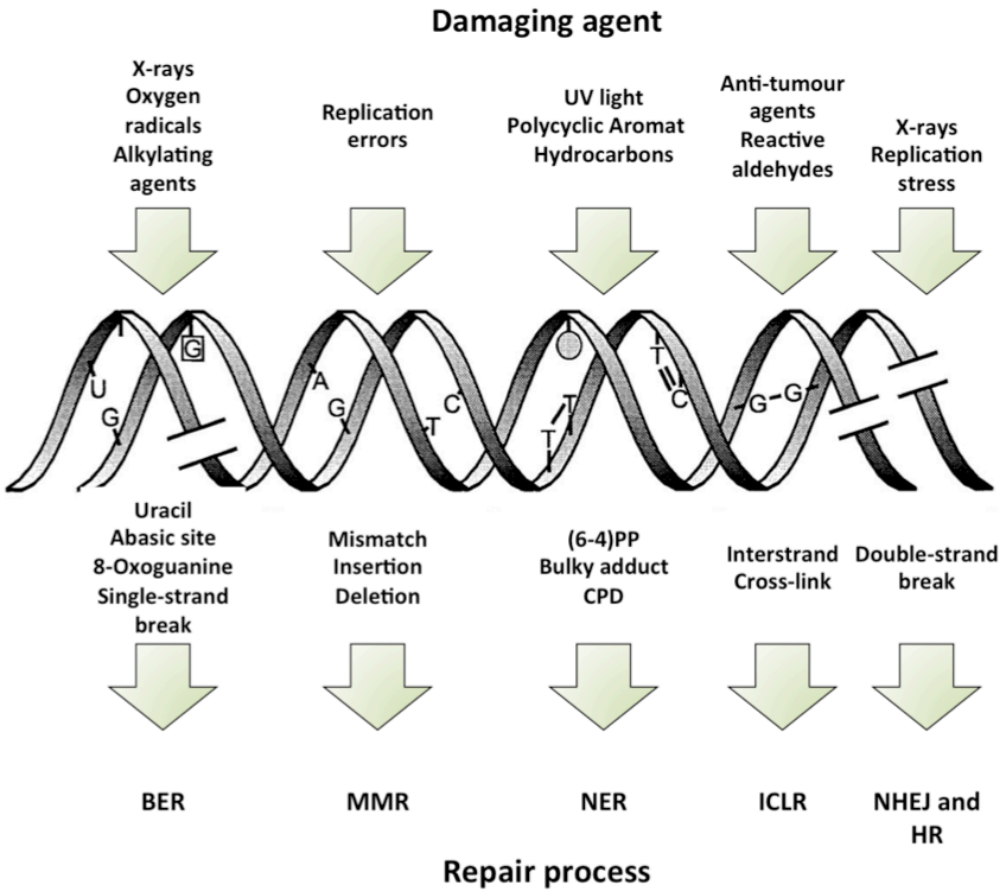

Figure 1: DNA damage and repair. An overview of the different DNA damaging agents, their effects on the DNA and repair pathways. Figure adapted from (1).

The cell has evolved many ways to counteract DNA damage and hence the formation of cancer. An important part consists of the DNA metabolic machineries that can repair DNA damage. There is a multitude of different DNA repair pathways, each specialized to deal with a particular type of damage (see **Figure 1**). These pathways include the repair of replicative

errors by proofreading of the polymerases (4) or by the replication-associated mismatch-repair (MMR) machinery (5), direct reversal of the damage (6), base excision repair (BER) for the repair of damaged or missing bases (7), nucleotide excision repair (NER) for removing bulky adducts (8), non-homologous end joining (NHEJ) (9) and homologous recombination (HR) (10,11) for the repair of double-strand breaks (DSBs) and finally, the repair of interstrand cross-links (ICLs) by a special pathway involving several mechanisms (12). The basic steps in each DNA repair pathway are the recognition of the lesion, cleavages that are made in the DNA to remove the damage, re-filling of the resulting gap by polymerases and finally ligation to restore the genetic information. These basic steps are heavily adapted to the specific needs of each DNA repair pathway.

It has to be mentioned, however, that in special cases, DNA damage can also be tolerated (13). These DNA damage tolerance (DDT) mechanisms include translesion synthesis (TLS) and template switching (14).

1.2 Replication and replication stress

1.2.1 Eukaryotic DNA replication

The faithful replication of the genetic material during S (synthesis) phase of the cell cycle is a prerequisite for the cell division in M (mitosis) phase. Eukaryotic DNA replication originates from thousands of origins of replication in each cell division and proceeds in a bidirectional manner. Interestingly, not all origins of replication are activated at the same time and some of them are 'dormant', meaning that they are never used under normal conditions, but can be activated during stress. Origins of replication are activated in a three-step process, including the recognition of origins, the assembly of a pre-replication complex (in G (gap) 1 phase of the cell cycle) and activation of the pre-replication complex. This process is highly regulated through the cell cycle to prevent multiple rounds of replication during a single cycle. Due to the antiparallel nature of DNA (one strand being 5'->3' and the other one 3'->5') and the fact that DNA polymerases always synthesize DNA 5'->3', the two strands must be replicated in opposite directions at the replication fork (see **Figure 2**). As the replication machinery is coordinated in a way that both strands are synthesized at the same site, one strand, the leading strand, can be synthesized continuously, while the so-called lagging strand must be synthesized in a discontinuous manner, producing stretches of DNA of about 100 to 200 bases, called Okazaki fragments. Owing to the asymmetric nature of DNA synthesis, the DNA has to be unwound ahead for the whole length of at least one Okazaki fragment before it can be used as a template for synthesis. This leads to a long stretch of single-stranded DNA (ssDNA) that is coated by the heterotrimeric complex replication protein A (RPA), important for the stabilization of the ssDNA and the prevention of secondary structures. An active replisome further contains the DNA polymerases Pol ε and Pol δ, which perform leading and lagging strand synthesis, respectively. However, DNA polymerases require a primer to begin synthesis and Pol α first acts as a primase, synthesizing a short stretch of RNA and about 10

to 20 deoxynucleotides (15). The RNA primer is later removed in a process called Okazaki fragment maturation. During replication, Pol δ displaces a stretch of the previous Okazaki fragment at each end of synthesis, creating a RNA-DNA single-stranded flap. Short flaps are cleaved off by FEN1 (flap endonuclease 1), a structure-specific nuclease, resulting in a nick, which can be sealed by Ligase I. The processing of longer flaps involves the action of DNA2 (DNA replication ATP-dependent helicase/nuclease), which is able to cleave longer flaps that are covered with RPA and therefore refractory to FEN1 cleavage. DNA2 action leaves behind a shorter flap that can again be processed by FEN1. Recently, it has been shown that yeast Dna2 can by itself generate re-ligatable ends and cleaves off the flap completely *in vitro* (16). Furthermore, the replisome contains replication factors, one of which being RFC (replication factor C, clamp loader), a five-subunit complex responsible for loading proliferating cell nuclear antigen (PCNA) on the DNA, which in turn is important for the stable interaction of DNA polymerases with the template DNA and their processivity (15). The replicative helicase consists of the Cdc45-MCM-GINS complex (CMG) and unwinds the DNA in front of the DNA polymerases by the ATP-driven 3'->5' helicase activity of the MCM2-7 hexamer (mini-chromosome maintenance). CMG also has a central role in coupling the helicase activity with the leading and lagging strand polymerases, as 'running-off' of the helicase would lead to long stretches of vulnerable ssDNA. Furthermore, the replisome also associates with checkpoint proteins in order to survey the DNA replication accuracy (see below) (17).

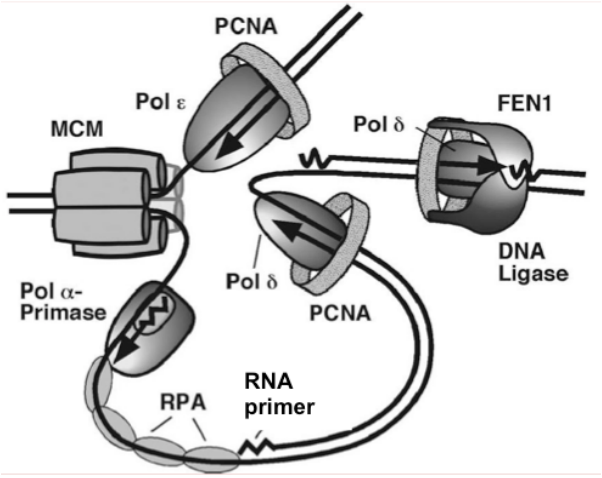

Figure 2: The eukaryotic replisome complex. A schematic representation of the eukaryotic replication fork is shown including only the most important proteins. Figure adapted from (18).

Pol δ and ε both contain a 3' exonuclease activity used for proofreading. The correct nucleotide is bound rapidly in the active center of the polymerase and the correctness is assessed by a conformational change in which tight contacts are made between the active site and the newly forming base pair. Only if the bound nucleotide is correct is the primer extended, otherwise the terminal mismatch hinders extension and the polymerase switches to the exonuclease complex to remove the wrong nucleotide. Nucleotide incorporation then resumes at the trimmed end. The replicative polymerases thus repetitively shuttle between a

polymerizing and editing mode, which is regulated by the (in)correctness of the incorporated nucleotide (19).

1.2.2 Replication stress

Replication stress can be defined as the slowing or stalling of replication forks and DNA synthesis. The sources and consequences of replication stress are very diverse, which makes it difficult to unambiguously characterize this condition. Sources of replication stress include nicks or ssDNA gaps that are left in the DNA as intermediates of DNA repair pathways. They can also result from topological release or arise spontaneously (20). Another important source of DNA replication stress are unrepaired DNA lesions, including abasic sites, pyrimidine dimers, alkylated bases, bulky adducts and cross-links. Such unrepaired lesions can be bypassed by the DDT pathway and be repaired during postreplicative repair (21). Furthermore, misincorporated ribonucleotides can cause replication stress, which are inserted into the DNA by the replicative polymerases Pol δ and Pol ε at a strikingly high rate. Those ribonucleotides can be removed from the DNA by RNase H2, together with FEN1 or EXO1 (exonuclease 1) (22). Certain sequences, like trinucleotide repeats, repetitive sequences or those that can form secondary structures and G-quadruplexes are also hard to replicate and need to be resolved by helicases for efficient replication. The collision of transcription bubbles with replication forks is another issue for the replisome. The cell employs different strategies to overcome this problem, e.g. the release of transcribed mRNA to stop transcription when a replication fork arrives. Another important mechanism is the covering of new transcripts by RNA processing factors, which prevent the formation of RNA-DNA hybrids, the so-called R-loops, that can also present a problem for replication. Topoisomerases and helicases can release the topological stress that arises through the vicinity of replication and transcription machineries and RNase H as well as RNA-DNA helicases can remove R-loops. Common fragile sites are regions in the genome that are prone to replication stress-induced DSBs, which might be due to a lack of replication origins in this region and the inability to rescue forks. Surprisingly, the breaks at such common fragile sites are not induced passively, but are a product of MUS81-catalyzed cleavage (23,24). In addition, when factors important for DNA replication are limited, this can slow down replication speed and lead to the accumulation of replication stress. Limitation in the nucleotide pool is for example an important source of replication stress and of cellular transformation (25). Replication stress can be induced by hydroxyurea (HU), which inhibits ribonucleotide reductase, leading to a decreased production of deoxyribonucleotides. Aphidicolin (APH) is another drug commonly used to actively induce replication stress, as it inhibits the replicative polymerases.

The persistence of ssDNA, bound by RPA and adjacent to double stranded DNA (dsDNA), also initiates the replication-stress response. The role of this signaling cascade is to arrest the cell cycle to allow time for repair and prevent the cell from entering mitosis. The so-called checkpoint factors also directly stabilize replication fork structures and facilitate DNA repair. Checkpoint signaling is activated after DNA damage or replication stress. Two main kinases

initiate this complex network: ATM (Ataxia telangiectasia mutated) and ATR (Ataxia telangiectasia and Rad3 related), which belong to the PIKK (phosphatidylinositol-3-kinase related kinases) family. These two kinases preferentially phosphorylate their targets at SQ/ST motifs (26). ATM is mainly responsible for the checkpoint response after DSBs, while ATR, together with its partner ATRIP (ATR-interacting protein), is activated through stretches of ssDNA covered by RPA. However, as ssDNA always exists at unperturbed replication forks on the lagging strand, it is important that ATR distinguishes those from stressed replication forks. This is mainly done by the 9-1-1 (Rad9-Hus1-Rad1) heterotrimeric clamp and its loader RFC^{Rad17}. They recognize the damaged DNA and facilitate the interaction between ATR-ATRIP and checkpoint mediators (i.e. TopBP1 and Mrc1/Claspin), leading to the phosphorylation of checkpoint kinase 1 (Chk1), the major downstream checkpoint effector kinase (15).

1.2.3 Recovery of stalled replication forks

There are different ways how a damaged or stalled replication fork can resume synthesis (27). The easiest way to restart replication beyond a lesion is in the case of a lagging-strand template lesion, where the next Okazaki fragment can simply reprime downstream of the damage. However, even on the leading strand it is possible to continue DNA synthesis downstream of the damage by repriming, leaving behind a ssDNA gap that is repaired postreplicatively (28). This means that both, leading and lagging strand synthesis can be discontinuous. Alternatively, dormant origins can be activated so that replication is completed by the new replication fork arriving from the other side. Another way to continue DNA synthesis downstream of the lesion without leaving a ssDNA gap behind is by using TLS polymerases, which are able to bypass the damage so that accurate synthesis can continue. Indeed, also in the case of replication fork recovery by repriming, the remaining stretch of ssDNA can be filled-in by TLS polymerases or by template switching.

One of the major sources of genome instability in eukaryotic cells is the collapse of stalled replication forks, leading to the formation of DSBs. If lagging strand synthesis continues once the leading strand is blocked at the site of a lesion, this can result in the uncoupling of the replicative helicase from leading strand synthesis and hence a large stretch of ssDNA will be left behind. This ssDNA is vulnerable and can lead to fork collapse, either passively or actively by endo- or exonuclease mediated cleavage (29). The breaks resulting from replication fork collapse are repaired by HR (see below, **chapter 1.3.6**), allowing for recombination-mediated fork restart.

If direct replication fork restart by the activation of dormant origins or repriming fails, there is another way of rescuing stalled replication forks without the need for recombinational repair. This pathway involves complex replication fork remodeling by the regression of the replication fork and re-annealing of the template as well as the newly synthesized strands (30). These reversed forks can be restarted by exonucleolytic degradation by DNA2, reverse branch

migration catalyzed by the annealing helicase SMARCAL1 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1), or by HR (31,32). The so-called chicken-foot structure that is formed during fork reversal allows the lesion to be repaired by NER, as it is returned to a region of dsDNA. Alternatively, the newly synthesized lagging strand that now anneals with the nascent leading strand can act as a template for DNA synthesis across the site of damage as during template switching. In that case, after reverse branch migration by SMARCAL1, the lesion has to be repaired in a post-replicative manner.

1.3 Interstrand cross-link repair (ICL repair)

1.3.1 Overview of the ICL repair pathway

ICLs are covalent linkages between the two DNA strands and are very cytotoxic as they impede almost all DNA metabolic processes. The repair of ICLs requires the action of multiple DNA repair pathways (see **Figure 3**). The coordination of repair happens through the so-called *Fanconi anemia* (FA) pathway, which recognizes the lesion during S-phase when the replication fork(s) are blocked at the site of the lesion. There is, however evidence for S-phase independent ICL repair in mammalian cells (see **chapter 1.3.10**). Incisions around the lesion allow for ICL unhooking, releasing it from one of the two strands and results in the formation of a DSB due to the present stalled replication fork. TLS restores the DNA template, which can be used during HR to complete the process by repairing the DSB. The remaining adduct coming from the unhooked ICL may eventually be repaired by NER. In replication-independent ICL repair, NER and TLS pols play a major role. Here, I shall discuss the complex mechanism of ICL repair, starting with a description of ICLs and the drugs generating them. I shall then introduce each of the mentioned processes in more detail and finally elaborate on ICL repair in different contexts. My focus will be on mammalian ICL repair and specifically on the protein I've been working with, FAN1 (FANCD2-associated nuclease 1).

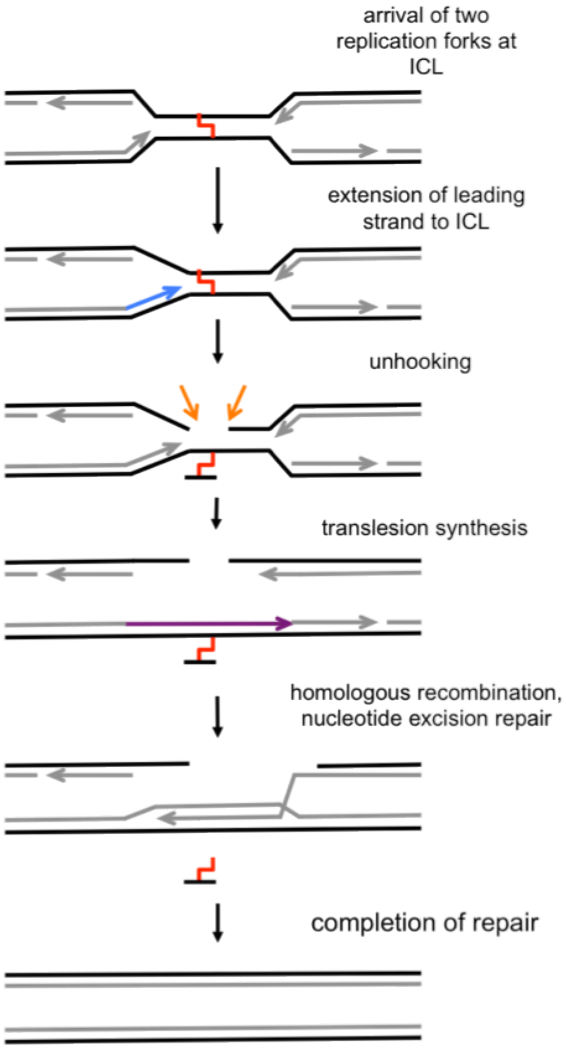

Figure 3: Overview of the ICL repair pathway. After stalling of one or maybe two replication forks at the ICL (red), one is extended (blue) to the site of the ICL for nucleolytic incisions (orange arrows) to be made in the unhooking step. TLS polymerases refill the resulting gap (purple) and HR is used to repair the DSB formed during unhooking. NER probably removes the remaining adduct and the DNA repair process is completed.

1.3.2 Types of ICL inducing agents

ICL agents can be of exogenous or endogenous origins and in general are alkylating agents with two reactive groups. One difficulty in studying the complex repair pathway that takes care of ICLs comes from the vast variety of ICLs formed by diverse agents that can elicit distinctive responses in the cell. Outcomes like sensitivity, cell cycle arrest or repair pathway choice can vary depending on the drug used (33). Therefore, and together with the fact that pathway choice and response also depend on the model system used, it is almost impossible to define a single ICL repair pathway and one has to be aware that sometimes a particular finding cannot be applied in a general context.

Mustard gas (or sulphur mustard) reached tragic popularity during the Second World War, when a secret cargo of mustard gas bombs was destroyed at the air raid on Bari. Several hundred soldiers and civilians were affected and examination of the survivors showed a decreased number of lymphocytes. Exposure to mustard gas often resulted in slow and painful death, causing severe skin oedema and ulceration, blindness and suffocation. Unexpectedly, research with this compound resulted in the first chemotherapeutic agent.

Clinically used nitrogen mustards (NMs; derivatives of mustard gas) include cyclophosphamide and chlorambucil for the treatment of leukemia, melphalan for myeloma and ifosfamide for the treatment of testicular, breast and lung cancer (34). About 5% of the DNA damage caused by (NMs) are ICLs, while monoadducts and protein DNA cross-links constitute the majority of the products. The ICLs are mainly formed between the N7 atoms of guanines within GNC sequences. NM-derived ICLs are highly DNA distorting and relatively unstable (33).

Other ICL agents include nitrosoureas (like BCNU, CCNU or methyl-CCNU), platinum compounds (like cisplatin), diepoxybutane (DEB), mitomycin C (MMC), or psoralens that are activated by UVA light. MMC and psoralens are both natural-occurring sources of ICL agents (33,35).

ICL agents can also be produced endogenously in the cell. Aldehydes like malondialdehyde, formaldehyde, acetaldehyde, unsaturated aldehydes (acrolein, crotonaldehydes) or nitrous acid are all endogenous sources of ICL damage (12,35). Aldehydes are byproducts of metabolic pathways, like lipid peroxidation and alcohol metabolism and are found in certain foods or tobacco smoke. Nitrous acid is a byproduct of the metabolism of nitrite-containing food preservatives (34).

DNA containing synthetic ICLs is an important tool to study those lesions *in vitro.* First attempts to obtain cross-linked DNA were based on the treatment of DNA with ICL inducing agents and purification of the desired products. However, as these reactions usually yield more intrastrand cross-links and monoadducts and only a minor fraction consists of ICLs, this approach was not very efficient. The new strategies for the synthesis of cross-linked DNA include the cross-linking of two nucleosides outside DNA before extending them into an oligonucleotide or the introduction of ICL precursors into two separate DNA molecules, which can, after annealing of the two strands, be activated to form the ICL (35). One such strategy is to incorporate 7-deazaguanine residues bearing acetaldehyde groups on opposing DNA strands and the induction of major groove NM-like cross-links by double reductive amination (36).

1.3.3 *Fanconi anemia* proteins and pathway

Fanconi anemia (*FA)* is a rare genetic disease characterized by bone marrow failure, congenital abnormalities, predisposition to solid and hematological tumors and hypersensitivity to cross-linking agents (37). The hypersensitivity towards DEB, with a dramatic increase in chromosome aberrations and quadradials is hence widely used as a diagnostic test for FA. FA is caused by mutations in one of at least 19 different genes, forming the FA complementation group (FANC-A to FANC-T), which all function in the repair of ICLs (see **Figure 4** and **Table 1**) (38-40). The most commonly mutated gene in patients is FANCA, followed by FANCC and FANCG (38).

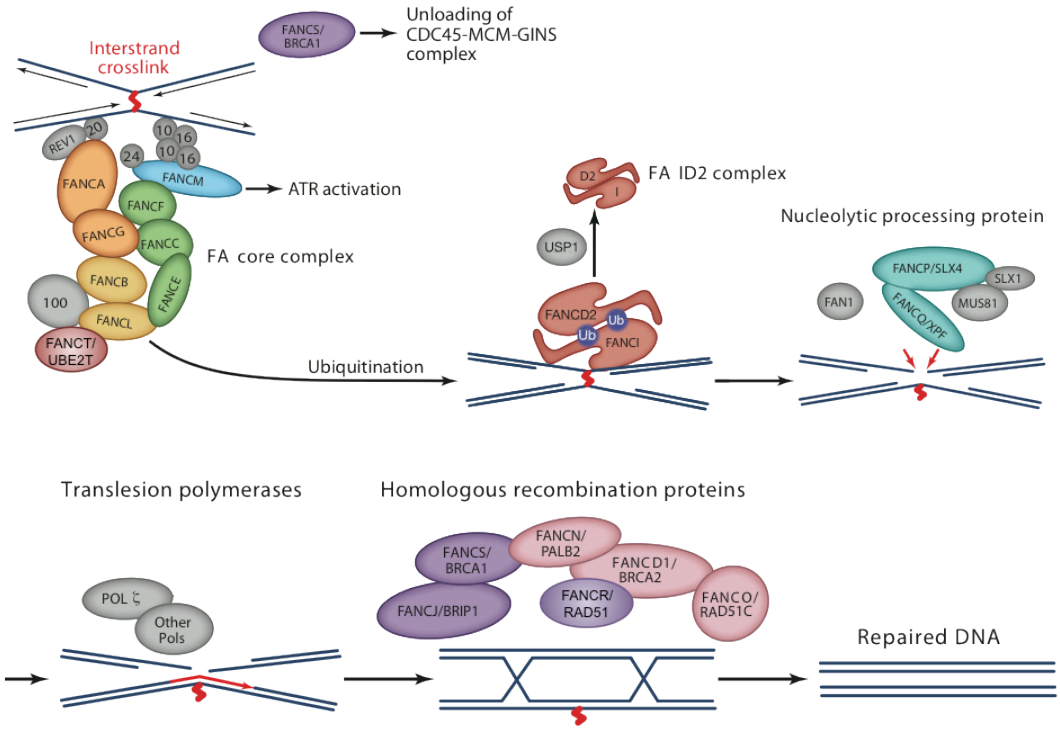

Figure 4: The FA pathway. The FA pathway starts with the recruitment of the FA core complex by FANCM and the unloading of the CMG complex. Ubiquitination of the FA ID2 complex then leads to nucleolytic incisions to unhook the ICL. TLS pols then re-fill the gap and HR repairs the DSBs that resulted from the unhooking step (FA effector proteins). FANC proteins are depicted in colors, while associated proteins, for which no FA patient mutations have been found, are shown in grey. Figure adapted from (39).

After the formation of ICLs, DNA replication stalls at the site of the lesion during S phase of the cell cycle. The FA pathway is then initiated by FANCS/BRCA1 (breast cancer 1), which, together with its partner BARD1 (BRCA1 associated RING domain), unloads the replicative CMG helicase from the stalled replication fork (41). However, patients with FANCS/BRCA1 mutations do not show bone marrow failure and therefore suffer from the so-called FA-like syndrome. FANCS/BRCA1 also has a much better known function downstream in the ICL repair pathway, namely in HR (see below, **chapter 1.3.6**). ICL recognition then requires FANCM (a DNA translocase and ATPase with branch migration activity), together with its partners FAAP24 (FA-associated protein 24kDa), MHF1 and MHF2 (histone fold protein 1 and 2), which are members of the FA associated proteins (42-44). This leads to the recruitment and assembly of the FA core complex (FANCA, B, C, E, F, G, L, M and T) (42,45). FANCM has additional functions in the activation of cell cycle checkpoints through ATR-Chk1 signaling (46). This signaling cascade also leads to the phosphorylation of multiple FA core complex factors, including FANCA, E, D2, and I (47). FANCM associates with chromatin throughout the cell cycle and it contains, in addition to its helicase domain, a degenerate ERCC4-like nuclease domain, which is required for the interaction with FAAP24. After recruitment of the core complex to the site of the cross-link, FANCL (a RING domain-containing E3 ubiquitin ligase), together with FANCT/UBE2T (an ubiquitin-conjugating

enzyme) (40), monoubiquitinates the FANCD2-FANCI complex (ID2) on both subunits (39). ID2 complex ubiquitination presumably stabilizes the complex binding to chromatin and is important for the coordination of the downstream steps, like nucleolytic incisions and translesion synthesis (48). USP1 is the responsible deubiquitinase for ID2 and this step is also essential for ICL repair. The ubiquitinated form of the ID2 complex acts as a landing pad for FAN1 and FANCP/SLX4, which both contain UBZ4 (ubiquitin-binding zinc finger 4) domains that specifically recognize the ubiquitin moiety of FANCD2 (47). FANCP/SLX4 in turn is a scaffold protein, recruiting several structure-specific nucleases, such as SLX1, MUS81-EME1 and XPF-ERCC1 (49). One or several of these endonucleases then perform the unhooking reaction. Unhooking leads to a suitable substrate for TLS polymerases to repair one strand but so far no mutation of a TLS pol gene has been associated with FA. HR repairs the remaining DSB coming from initial incisions. Indeed, the other FA proteins are all implicated in HR (i.e. FANCD1/BRCA2, FANCJ/BRIP1, FANCN/PALB2, FANCO/RAD51C, FANCR/RAD51 and FANCS/BRCA1) by facilitating RAD51 loading or the resolution of recombination intermediates.

Name	Function/ Activity	Incidence in FA patients (%)	Ortholog in *S. cerevisiae*?
FANCA	Core complex member	66	no
FANCB	Core complex member	0.8	no
FANCC	Core complex member	9.5	no
FANCD1/ BRCA2	Loading of RAD51 on ssDNA	3.3	no
FANCD2	DNA binding, DNA damage signaling and recruitment of repair factors, forms complex with FANCI	3.3	no
FANCE	Core complex member	2.5	no
FANCF	Core complex member	2	no
FANCG	Core complex member	8.7	no
FANCI	DNA binding, DNA damage signaling and recruitment of repair factor, forms complex with FANCD2	2	no
FANCJ	3'-5' helicase, acts during HR in association with BRCA1	1.6	yes
FANCL	Core complex member, E3 ubiquitin ligase that monoubiquitinates FANCD and FANCI	<1	no
FANCM	DNA translocase, recruits core complex to ICL, involved in checkpoint activation	<1	yes
FANCN/ PALB2	Assists BRCA2 in RAD51 loading	<1	no
FANCO/ RAD51C	Involved in HR during ICLR, DNA-dependent ATPase, involved in HJ resolution	<1	yes
FANCP/ SLX4	Scaffold for endonucleases	<1	yes
FANCQ/ XPF	3' flap endonuclease, involved in unhooking	<1	yes
FANCR/ RAD51	ssDNA binding for homology search during HR	<1	yes
FANCS/ BRCA1	Loading of RAD51 on ssDNA, pathway choice during HR	<1	no
FANCT/ UBE2T	E2 ubiquitin conjugating enzyme involved in FANCD2-FANCI monoubiquitination	<1	yes

Table 1: The Fanconi anemia complementation group.

1.3.4 Unhooking step in ICL repair

Even though unhooking does not remove the ICL completely, it allows bypass of the lesion by polymerases. Five different nucleases have been implicated in the unhooking step: XPF-ERCC1, MUS81-EME1, SLX1-SLX4, SNM1A and FAN1. FAN1 will be introduced separately in **chapter 1.5**.

General features of nucleases

Nucleases are molecular scissors that catalyze the cleavage of DNA or RNA between the sugar and the phosphate moieties of the backbone. They can be generally divided into exonucleases, which hydrolyze from the 5' or 3' end of nucleic acids, and endonucleases, which don't need a free end and hydrolyze internal phosphodiester bonds. However, the fundamental chemistry of endo- and exonucleases is the same and it is not uncommon that a single enzyme possesses both activities, mostly even using the same active site. The cleavage products of nucleases contain a 5' terminal phosphate group on one hand and a 3' terminal hydroxyl group on the other one. Virtually all DNA metabolic and repair pathways rely on DNA nucleases, with a few exceptions in the case of direct damage reversal. Moreover, many nucleases can act in more than one DNA repair pathway and a given cleavage reaction can usually be performed by multiple nucleases. This is what makes the study and understanding of these nucleases so challenging. DNA endonucleases can further be distinguished into sequence- damage- or structure-specific. Sequence-specific endonucleases are mainly known from prokaryotes, like restriction enzymes and the mismatch repair protein MutH. One example of eukaryotic sequence-specific endonucleases is I-SceI, which is expressed in the mitochondria of *S. cerevisiae* (50). Examples for damage-specific nucleases in mammalian cells are the apurinic/apyrimidinic (AP) endonucleases used in BER to cleave 5' from an abasic site (51). Structure-specific endonucleases recognize intermediates of DNA repair that have specific structures. An example are the NER enzymes XPF and XPG, which cleave at the base of a bubble structure that surrounds the DNA damage after unwinding. DNA nucleases are usually two-metal-ion-dependent, but there are also some that depend only on one metal ion, like the type II restriction enzyme nucleases, and some even do not need a metal ion for catalytic activity, like DNAse II (51). In most cases Mg^{2+} is the metal ion needed for catalysis and Ca^{2+} usually inhibits the reactions. Mn^{2+} can often also be used; it usually relaxes the specificity towards substrates and sometimes can even rescue defective enzymes (52). The most frequent use of Mg^{2+} could be due to its high abundance in living organisms, its solubility, redox stability, small size and rigid coordination geometry compared to other divalent metals (51). For the general reaction mechanism of two-metal-ion-dependent nucleases, see **Figure 5**.

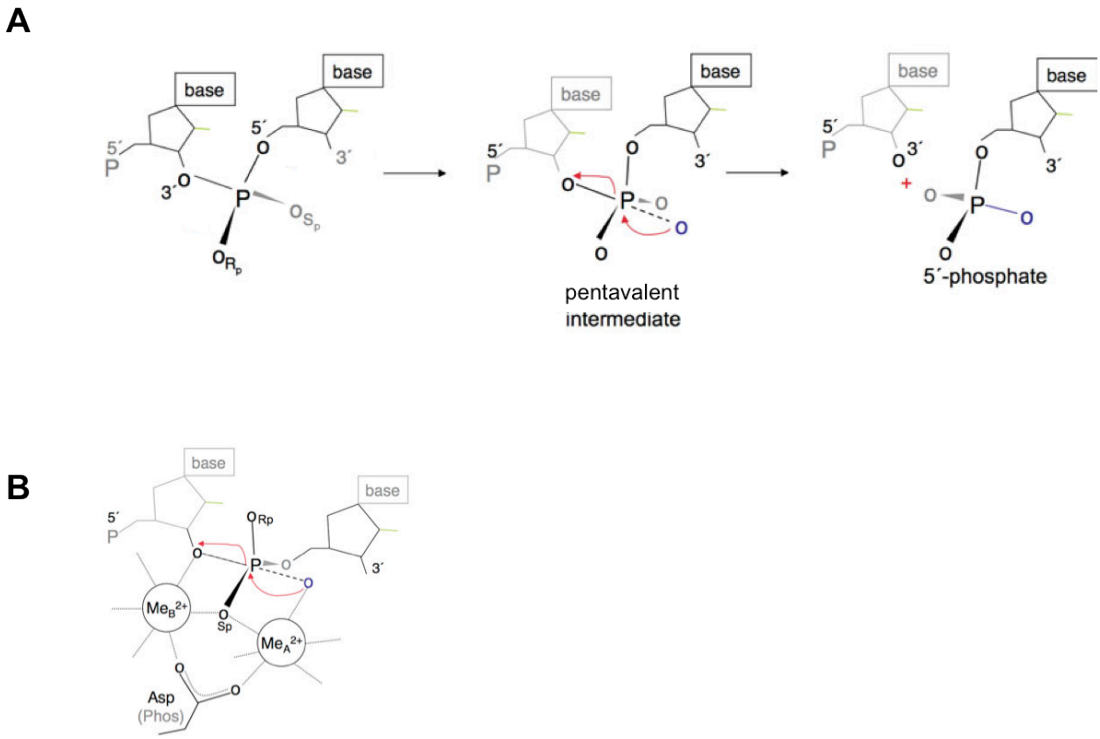

Figure 5: Reaction mechanism of nucleases. A) A nucleophilic attack by a deprotonated water molecule (the oxygen of which is shown in blue) leads to the pentavalent transition state. The P-O3' bond is then cleaved to generate a 5'-phosphate and a 3'-OH product. **B)** In the two-metal-ion-dependent enzymes, one of the oxygens coordinates both metal ions (Sp) and the A ion drives the reaction by deprotonating the nucleophilic water molecule (shown in blue). The B ion is used to stabilize the pentavalent transition state and to destabilize the ground state of the scissile phosphate for cleavage. Figure adapted from (51).

It is very difficult to classify nucleases, as their sequence, structure, reaction mechanism, or function usually don't correlate, which would result in different categories depending on the criteria used. Usually, nuclease families are defined according to the secondary structure of their nuclease domains or sequences of active-site motifs, as their function or tertiary structure are not always known.

XPF-ERCC1

XPF-ERCC1 (*Xeroderma pigmentosum* complementation group F- Excision repair cross-complementation group 1) is best known for its function in NER. The two subunits contain the same domains (a ERCC4 endonuclease domain and a Helix-hairpin-Helix (HhH) DNA binding domain) but the ERCC4 endonuclease domain is degenerate in ERCC1, rendering it catalytically inactive (53). XPF-ERCC1 preferentially cleaves splayed arm, bubble and stem-loop structures with a 3' flap polarity and cleavage occurs on dsDNA regions close to the ssDNA. It has a specific role in ICL repair independently of the NER pathway and cells deficient in XPF-ERCC1 are more sensitive to ICL inducing agents than cells defective for other NER proteins (54). Biochemically, XPF-ERCC1 was shown to be able to unhook psoralen ICLs *in vitro* (55). Indeed, XPF mutations were also found in *FA* patients, suggesting

that it works in an FA-dependent ICL repair pathway and giving it the alternative name FANCQ. XPF-depleted *Xenopus* egg extracts are not able to perform the initial incision, further indicating a role for this protein in unhooking. Interestingly, in this system both cleavages 5' and 3' of the ICL were defective, indicating that the first incision by XPF is needed for the second one to occur, or that XPF is itself able to perform both incisions. The unhooking activity of XPF-ERCC1 was further stimulated by SLX4 (56). In mammalian cells, it was shown that ICL-dependent DSBs still form and even persist in the absence of XPF-ERCC1 (54,57). This would be contradictory to a function in ICL unhooking. However, it might be reconciled with the fact that those stalled forks persisting in the absence of XPF-ERCC1 are aberrantly cleaved by MUS81 (58). Furthermore, it is unclear how the 3' flap endonuclease activity of XPF-ERCC1 might perform both cleavages in unhooking. Therefore, although XPF-ERCC1 is commonly accepted to be the main ICL unhooking enzyme, further studies are needed to elucidate its exact mode of action.

MUS81-EME1

MUS81-EME1 (Methansulfonate and Ultraviolet-Sensitive 81- Essential Meiotic structure-specific Endonuclease 1) belongs to the same nuclease family as XPF, with MUS81 being the catalytically active subunit. MUS81 is also a 3' flap endonuclease, but, in contrast to XPF, it prefers substrates where there is a 5' end close to the dsDNA-ssDNA junction (53). MUS81-EME1 has been implicated in ICL unhooking due to its importance for ICL resistance and the DSB reduction upon depletion of the protein (58,59). However, those cells are less sensitive to ICL agents than XPF-ERCC1 lacking cells and in the *Xenopus* egg extract system depletion of MUS81-EME1 did not have an effect on ICL unhooking (56). This suggests that it acts only in a subset of cases, potentially when XPF-ERCC1 cleavage cannot occur due to a 5' end close to the junction of the stalled replication fork(s).

SLX1-SLX4

SLX1 (structure-specific endonuclease subunit SLX1) contains a UvrC-intron-endonuclease domain (URI) and a PHD-type zinc finger domain and its preferred substrate is a 5' flap, which it cuts at the flap junction (60). SLX1's activity is drastically increased upon interaction with SLX4. SLX1 was proposed to be the main 5' flap endonuclease required for unhooking, as deletion of the gene causes ICL sensitivity, similar to that of MUS81 (61).

SLX4 (structure-specific endonuclease subunit SLX4) is a *Fanconi anemia* protein (FANCP) and deletion of SLX4 causes much greater sensitivity to ICLs than the depletion of SLX1 or MUS81 (56,61,62). It is also required for incisions in the *Xenopus* egg extract system (56). This suggests that SLX4 is an important scaffolding protein, coordinating the different nucleases - XPF, MUS81 and SLX1 - to the site of the cross-link. Interactions of SLX4 with XPF-ERCC1 and SLX1 are important for ICL repair, while the interaction with MUS81 is dispensable (61,62). As mentioned above, SLX4 contains a UBZ domain that is important for its recruitment to monoubiquitinated FANCD2 at the site of the ICL (49).

SNM1A

In vertebrates there are three proteins related to yeast Pso2/Snm1 in vertebrates: SNM1A (Sensitive to Nitrogen Mustard 1A), SNM1B/Apollo, and SNM1C/Artemis. They belong to the β-CASP subfamily of metallo-β-lactamases, which are all DNA processing enzymes (63). SNM1A is structurally most similar to Pso2, which made it the primary candidate for the vertebrate ortholog of Pso2, an endo- and exonuclease that has a specific function in ICL repair. SNM1C plays an important role in the DSB repair subpathway NHEJ and has so far not been implicated in ICL repair. SNM1B, which has functions in checkpoint signaling and telomere maintenance, could potentially act in ICL repair, as cells depleted of it are sensitive to ICL inducing agents (64,65). Human SNM1A and SNM1B are 5'->3' exonucleases, while SNM1C is a structure-specific endonuclease (66,67). SNM1A has a stronger affinity for single-stranded DNA over double-stranded DNA, which is not the case for SNM1B. SNM1A is also more active on high molecular weight DNA compared to SNM1B and even though both proteins are able to digest DNA past an ICL that is induced by the agent SJG-136, SNM1A has a greater capacity to do so (67). Therefore, its proposed role in ICL repair is to trim back the DNA around the ICL that might be remaining after unhooking and it seems to do so in a common pathway with XPF-ERCC1 (58). As predicted, SNM1A is able to rescue the sensitivity in Pso2 yeast mutants, which was not the case for SNM1B or SNM1C (68). This suggests that SNM1A is the functional ortholog of yeast Pso2. However, SNM1B might have a partially redundant function, as double disruption of the two genes is additive in terms of cisplatin sensitivity compared to the single disruptants in DT40 chicken cells (69).

SNM1A is recruited to the site of the ICL by RAD18-dependent ubiquitinated PCNA via its UBZ domain and PIP (PCNA interacting peptide) box (70). Just recently, the crystal structures of SNM1A and SNM1B/Apollo were resolved and a region of positive charge was found to surround the active site of SNM1A, which is absent in SNM1B/Apollo. This, together with a putative wide DNA-binding groove helps to explain the ability of SNM1A to accommodate and process highly distorted DNA structures (71).

1.3.5 Translesion polymerases and their roles in ICL repair

Translesion polymerases (TLS pols) are part of the DDT mechanism, also known as DNA damage bypass or post-replicative repair, where DNA replication is completed before the lesion is repaired. The repair then happens later in a post-replicative manner. There are two main ways of circumventing damage during replication, one are the TLS pols and the other one is template switching, in which the stalled nascent strand switches temporarily and uses the newly-synthesized strand of the sister chromatid as a template for synthesis past the lesion. Once it switches back to its original template, the newly synthesized strand is error-free. This is different in the case of specialized polymerases, which bypass the damaged base that would otherwise block the progression of the replication fork with the cost of being error-prone. TLS pols have different substrate specificities, enabling them to deal with

different types of damaged bases. The polymerase switch from replicative to TLS pols is mainly coordinated by PCNA ubiquitination upon fork stalling. The interaction between the ubiquitin moiety on K164 of PCNA and ubiquitin binding motifs on TLS pols ensures the recruitment of TLS pols to the stalled fork. Following the first polymerase switch, nucleotide insertion across the lesion and extension past the lesion may require the concerted action of more than one TLS pol and, once the lesion is bypassed, a further polymerase switch has to occur to reposition the replicative polymerase at the primer terminus and for accurate DNA synthesis to resume (72).

Based on sequence and structural similarities, human pols can be grouped into five different families: A, B, X and Y and the reverse transcriptase activity of telomerase. The common feature of all pols is that they fold into a structure that resembles a human right hand composed of three distinct domains, the palm, thumb, and fingers. In contrast to the replicative polymerases, TLS pols have a low fidelity in nucleotide incorporation and additionally lack the proofreading activity, making them error-prone. Whereas the fingers in replicative pols tightly bind the incoming dNTP, the active site of TLS pols is more open and less constrained to reject wrong base pairs. However, there are a few examples, as for Pol η with TT-CPDs (cyclobutane pyrimidine dimers), where the bypass of lesions can be error-free. The replicative pols α, δ and ε belong to the B family and the mitochondrial pol γ to family A.

Several TLS pols have been implicated in ICL repair, such as the Y family pols η, κ, and Rev1, the B family pol ζ, the A family pol ν and sometimes θ.

As mentioned before, Pol η, also called XPV, is responsible for the bypass of UV-induced CPDs. However, it also has important functions in ICL repair, highlighted by the fact that XPV patient-derived cell lines are highly sensitive to cross-linking agents (73,74). Pol η may also have a role in replication-independent ICL repair, as shown by a plasmid-based host-cell reactivation assay (75). Pol η is able to bypass synthetic, structurally distinct ICLs also *in vitro*, which becomes more efficient when using longer ICLs and a shorter DNA fragment surrounding the ICL opposite the template strand, both probably due to an easier displacement of the non-template strand (76).

Pol κ is the most accurate Y family polymerase on undamaged DNA and has a role in NER in the repair synthesis step together with Pol δ (77). Also Pol κ could potentially be involved in ICL repair, as human cells depleted of the protein are hypersensitive to MMC and it is able to fulfill ICL bypass *in vitro* (76,78).

Pol ζ is a heterodimeric protein, consisting of the catalytic subunit REV3 and the structural subunit REV7. Pol ζ is thought to be the major contributor to TLS over lesions. Its efficiency of lesion bypass is enhanced by REV1 and PCNA (79). Pol ζ has been implicated in ICL repair, as cells deficient for REV3 or REV7 are exquisitely sensitive to ICL inducing agents and in chicken cells the sensitivity was even greater than the one for FANCC (80). Also biochemically, (truncated) Pol ζ is able to bypass ICL lesions, in conjunction with Pol η (81). REV1 is not a real polymerase, having a template-directed deoxycytidyltransferase activity,

mainly incorporating C opposite DNA lesions. REV1 binds monoubiquitinated PCNA via its ubiquitin binding motif (UBM) and can interact with several TLS polymerases including Pol η, ι, κ and ζ via its C-terminus. In addition, FAAP20, a component of the FA core complex, was shown to be required for the recruitment of REV1 to nuclear foci and this happens through a direct interaction between a UBZ4 domain of FAAP20 and the monoubiquitinated form of REV1 (82). An additional role for REV1 together with Pol ζ appears in replication-independent ICL repair (83). Furthermore, in *X. laevis* extracts REV1 inserted a cytosine opposite the unhooked ICL, which was followed by Pol ζ extension (84).

Pol ν is also required for ICL repair, as evidenced by cellular sensitivity to cross-linking agents and its interaction with the ID2 complex. It further has a role in HR together with RAD51 and HELQ, a 3'->5' helicase (85). Pol ν is error-prone, has a robust strand displacement activity and the ability to bypass thymine glycols and major groove ICLs *in vitro* (86,87).

1.3.6 Homologous recombination and ICL repair

Double strand breaks (DSBs) are very cytotoxic lesions, which is mainly reflected by the fact that already the presence of only a few DSBs in the cell can lead to cell death and chromosomal rearrangements. The rearrangements in turn can potentially inactivate tumor suppressor genes or lead to the overexpression of oncogenes thereby driving cancer formation. Therefore accurate and fast repair of such damage is crucial to the cell and three main pathways have evolved to take care of these lesions in mammalian cells: Non-homologous end-joining (NHEJ), single-strand annealing (SSA) and homologous recombination (HR). HR is the most accurate mechanism, where the homologous sequence of the sister chromatid is used as a template for DNA synthesis, thereby restoring the missing DNA information. HR is however limited to the S and G2 phases of the cell cycle, as sister chromatids are only present then. In NHEJ on the other hand, the two termini of the broken DNA are processed, recessed and directly re-ligated, which can result in loss or gain of nucleotides, hence making it an error-prone mechanism of DSB repair. Moreover, if more than two free DNA ends exist in the cells, the wrong ones can be joined together, leading to chromosomal translocations. However, as NHEJ does not depend on homologous sequences, this pathway is preferably used during G1 and M phases of the cell cycle (88). The third pathway, SSA is even more inaccurate as it can lead to large deletions in the genome. It depends on direct repeats that are located close to the DSB, where extensive resection occurs in order for the two homologous sequences to anneal.

Besides their main role in the repair of DSBs, HR proteins have many other functions, like in the repair of DNA cross-links, at damaged replication forks, during the segregation of homologous chromosomes in meiosis I and in telomere maintenance (89-91). Several genetic disorders are associated with HR dysfunction, such as Bloom's syndrom, *Fanconi anemia* and cancer types including breast, ovarian and others. The HR pathway is schematically depicted in **Figure 6**. A prerequisite for homologous DNA recognition is the formation of

ssDNA by DSB end resection. In human cells, resection is initiated by the MRN complex, consisting of MRE11, RAD50 and NBS1. MRN is a Mn^{2+}-dependent exonuclease that resects DNA 3'->5' and is regulated by CtIP (CtBP interacting protein) (92). As HR requires 3' overhangs in order to start invasion of the complementary sequence, the MRN polarity seems to be incompatible with this function. However, it was shown in yeast that the ortholog of MRN, MRX (Mre11-Rad50-Xrs2) contains an endonuclease activity, which requires activation by the CtIP ortholog Sae2 and Mg^{2+} (93). It can therefore incise the dsDNA distal to the DSB and then degrade it 3'->5' until the DNA end to leave a 3' tail on the complementary strand (94). It is very likely that a similar mechanism exists in human cells, where CtIP could stimulate MRN endonuclease activity. MRN also has a role in sensing and signaling the DSB, leading to the activation of ATM, which in turn phosphorylates the histone variant H2AX on Serine 139 at the breakage site. This results in distinct γ-H2AX foci that are readily detectable by immunofluorescence (IF) (95). The short-range resection performed by MRN and CtIP is followed by a more extensive resection, executed by two different enzyme complexes. One is the BLM helicase together with the DNA2 nuclease. Although DNA2 also has helicase activity, this is dispensable for its function in DSB resection. RPA and the TOPOIIIα-RMI1-RMI2 complex both stimulate the helicase activity of BLM and ensure 5'->3' polarity of resection. The second pathway consists of EXO1 that is stimulated by BLM, MRN, RPA and CtIP (96,97). The 3' overhanging tails resulting from resection lead to the activation of ATR and are covered by RPA to eliminate secondary structures. In a next step, RAD51 is loaded onto the ssDNA to promote homologous DNA pairing and strand exchange in an ATP-dependent reaction. RAD51 binding to the DNA requires several proteins, including BRCA2, RAD51 paralogs, PALB2 and RAD52. During the so-called synapsis, the RAD51 filament invades the undamaged strand to generate a displacement loop (D-loop), where it primes DNA synthesis and leads to the formation of a Holliday junction (HJ). From here on there are several HR subpathways, mainly the synthesis-dependent strand annealing (SDSA) and dHJ (double Holliday junction) pathways. SDSA involves the re-annealing of the extended strand with the original complementary 3' tail, where DNA synthesis proceeds and any remaining flaps are cleaved and the nick is ligated. The main outcomes are non-crossover products, which reduces the potential for genomic rearrangements and makes it the preferred HR pathway (98). The dHJ pathway involves second-end capture and, as the name suggests, formation of a double Holliday junction as shown in **Figure 6**. Three main classes of eukaryotic structure-selective endonucleases that can cleave HJ have been identified (99). As for human proteins, the XPF class is represented by XPF and MUS81, the UvrC/URI-YIG class by SLX1 and the Rad2/XPG class by GEN1 (53,100). Depending on the cleavage polarity of the HJ resolvases, the result can be a non-crossover or a cross-over product. Alternatively, dHJs can be dissolved by the RecQ-family helicase BLM together with TOPOIIIα-RMI1-RMI2, leading to non-crossover products (98).

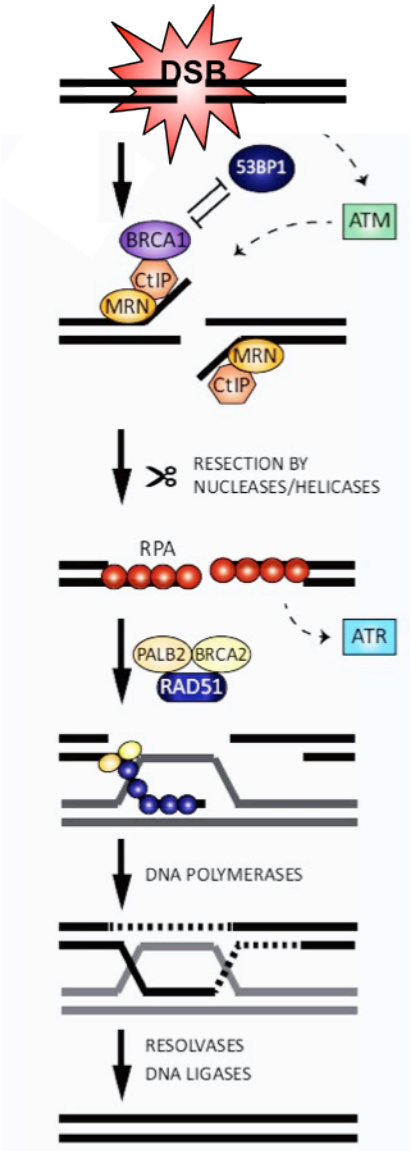

Figure 6: The HR pathway. After recognition of the DSB, pathway choice is directed towards HR by inhibiting 53BP1. The DNA end is resected 5'-3' by specialized nucleases, which leads to a 3' overhang. Strand invasion, D-loop formation and second end capture lead to the formation of a double HJ, which can be resolved and the DNA is re-ligated. Figure adapted from (101).

In the case of two replication forks stalling at the site of an ICL and unhooking (as shown in **Figure 3** for example), the result is a two-ended DSB, which can be repaired as just mentioned. If only one replication fork stalls at the cross-link and unhooking takes place, it results in a one-ended DSB, which can be repaired by the HR-subpathway called break-induced replication (BIR, shown in **Figure 7**). After resection of the free end, the 3' overhang can invade the other duplex and form a D-loop. As opposed to the other HR pathway, the newly synthesized strand does not have a complementary strand to re-anneal to. Therefore it continues replication until the end of the chromosome or the encounter with another replication fork. Lagging strand synthesis is re-established on the displaced DNA strand, resulting in a HJ that can be cleaved by HJ resolvases. Since BIR usually results in loss of heterozygosity (LOH), this pathway might be actively suppressed when the DSB contains two ends (102,103).

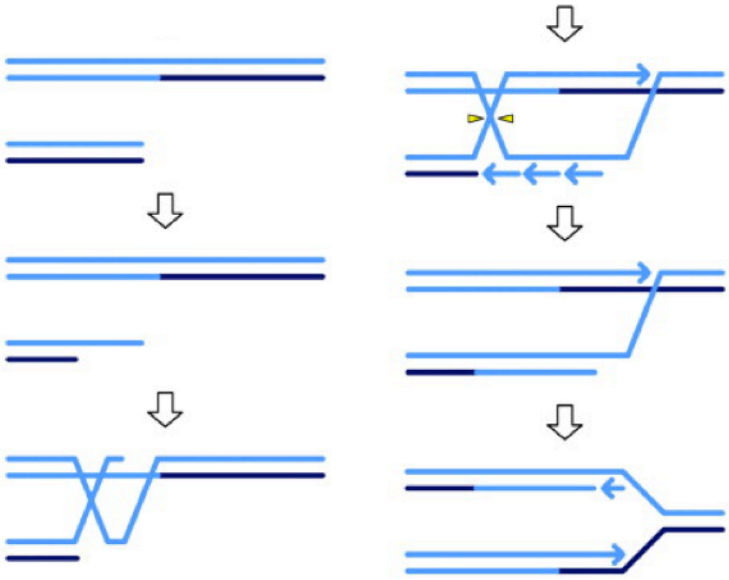

Figure 7: Repair of a one-ended DSB. Upon resection, the overhanging 3' tail invades the sister chromatid and creates a D-loop. Lagging strand synthesis resumes, which leads to the formation of a HJ that can be cleaved and the replication fork is restored. Figure adapted from (102).

As HR mediated repair is mainly restricted to the S and G2 cell cycle phases, the expression of many HR factors is elevated in these cell cycle phases (104). A further control occurs via phosphorylation of important players, especially those involved in the decision-making step of DNA end resection, by CDKs (cyclin dependent kinases) (105). For example, CtIP phosphorylation, stability and BRCA1 interaction are regulated during the cell cycle, to ensure the correct repair pathway choice (106-108). BRCA1 and 53BP1 are the two critical components that compete with each other for pathway choice, where 53BP1 promotes NHEJ and BRCA1 commits cells to HR (104,109). BRCA1 inhibits 53BP1 and therefore NHEJ, which is evidenced by the fact that BRCA1-deficient phenotypes (resulting from error-prone NHEJ) can be alleviated by 53BP1 deletion (110,111).

HR and ICL repair are tightly connected, which is mainly reflected by the fact that many of the FA proteins are actually known for their function in HR (BRCA2/FANCD1, PALB2/FANCN, BRIP1/FANCJ/BACH1, RAD51C/FANCO, RAD51/FANCR, BRCA1/FANCS). Furthermore, it was shown that patient-derived FANCA-, FANCG-, and FANCD2-deficient cells are slightly defective in HR repair (112) and that FANCC knockout DT40 cells are impaired in HR (113,114). On the other hand, the HR factor CtIP is recruited to sites of ICLs by FANCD2 and this complex plays an important role in the faithful repair of such lesions (115,116).

Interestingly, it has been shown that the inhibition of the NHEJ pathway in human FA patient-derived cell lines reduces the toxicity of ICL-inducing agents (117,118). The occurrence of radial chromosomes normally observed in FA cells was decreased upon inactivation of NHEJ pathway, indicating that NHEJ is responsible for these chromosomal aberrations and that the FA pathway is involved in the suppression of NHEJ during ICL repair to prevent the formation of such toxic genomic rearrangements. Furthermore, it shows that HR is still active in FA cells, but that the FA pathway either directly inhibits the NHEJ pathway or stabilizes a DNA structure that is more compatible with repair by HR than by NHEJ. Therefore, the toxicity

mainly comes from those genomic rearrangements caused by NHEJ and not from the inability to repair DSBs.

1.3.7 Nucleotide excision repair and ICL repair

NER is a highly versatile repair pathway, which takes care of different bulky, helix-distorting types of DNA lesions. UV lesions such as cyclobutane pyrimidine dimers (CPD) and 6-4 photoproducts are the most prominent modifications addressed by the NER pathway (119). NER additionally recognizes and repairs DNA lesions that are caused by chemicals that have the ability to covalently bind to one or two DNA bases, therefore forming bulky adducts and intra- or interstrand cross-links (120). The importance of NER is reflected by the severe genetic disorders that defects in NER cause, including *Xeroderma pigmentosum, Cockayne's syndrome* (CS) and *trichothiodystrophy* (TTD), all of which are characterized by an extreme sensitivity towards UV light (120). In addition, deficiency of the NER protein XPF has only recently been linked to *Fanconi anemia* (121,122).

Although not directly proven so far, it is thought that the remaining adduct on the DNA resulting from ICL unhooking is removed by nucleotide excision repair (NER). Furthermore, NER might have an important function in the removal of ICLs outside of S-phase (see below, **chapter 1.3.10**). XPF-ERCC1 depletion shows the highest sensitivity to ICL agents of the NER proteins, which is probably attributed to its independent role in ICL unhooking. However, also depletion of other NER proteins leads to sensitivity (54).

NER can generally be broken down into five steps: damage recognition to initiate repair, helix opening and unwinding, incision on either side of the lesion to release the damage-containing oligonucleotide, DNA synthesis and ligation (**Figure 8**). Furthermore, NER can be divided into two distinct subpathways, termed global genome repair (GGR) and transcription-coupled repair (TCR). As the names suggest, the former is responsible for the repair of lesions in the entire genome and the latter repairs transcription-blocking lesions present in transcribed genes. These two subpathways differ in the way of damage recognition. In GGR the XPC-HR23B (RAD23 homolog B) complex is the first NER factor to detect a lesion, whereas some less distorting lesions require initial recognition by the DDB1-DDB2 (damage-specific DNA binding protein 1 and 2) complex. In TCR, RNA polymerase II (RNAPII) is stalled by lesions in the transcribed strand of active genes and attracts NER enzymes, bypassing the need for XPC. CSA, CSB and XAB2 (XPA binding protein 2) are also specifically required for TCR. At this stage the two pathways converge, the TFIIH complex is recruited and immediately joined by XPA, RPA, and XPG. The XPB and XPD components of TFIIH are DNA helicases, and through their action TFIIH unwinds the DNA surrounding the lesion until a 30-nucleotide (nt) "bubble" is formed. RPA and XPA stabilize the DNA bubble and help to position two endonucleases at the bubble junctions. The first incision, on the 3′ side of the bubble relative to the lesion, is made by XPG, which also coordinates the second incision on the 5′ side of the bubble by the XPF protein and its partner ERCC1. A lesion-containing DNA fragment of

25–32 nts is released, the gap is filled in by DNA pol δ, ε or κ using the information from the intact complementary strand and the final ligation step can be carried out by ligase I and FEN1 or by the ligase III–XRCC1 complex (120,123,124).

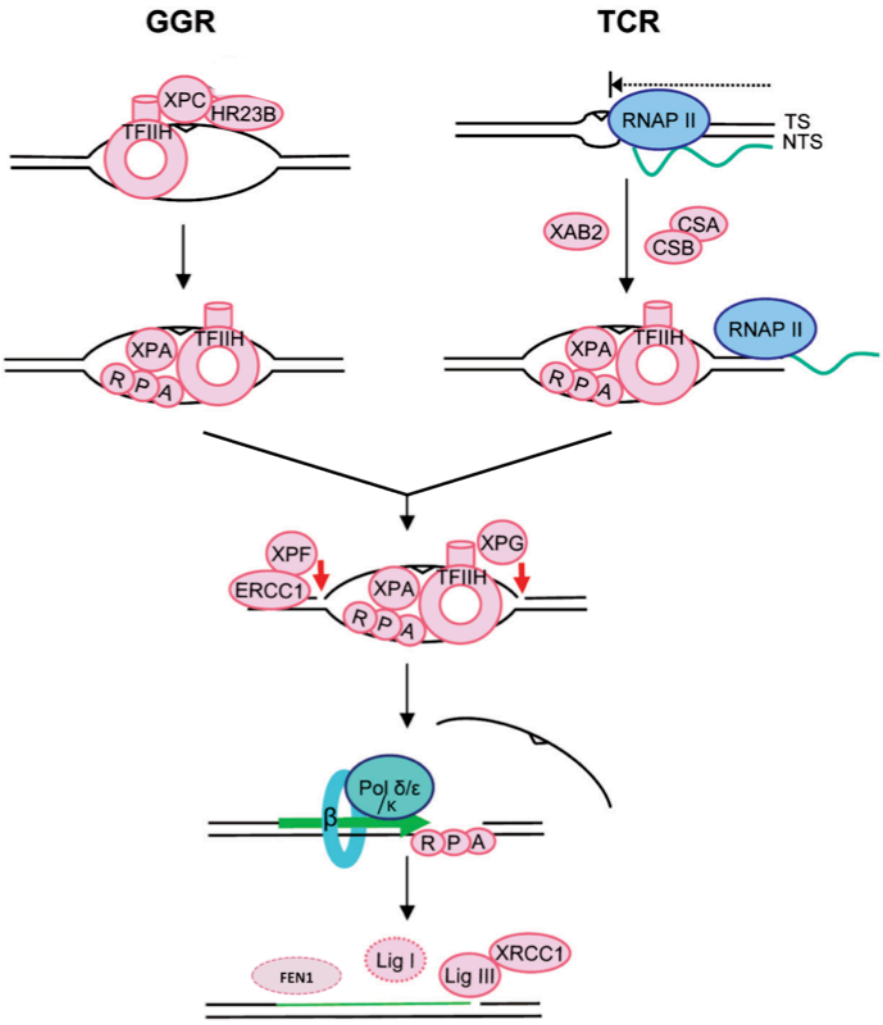

Figure 8: The two NER subpathways. The lesion (depicted as a triangle) is recognized by XPC and HR23B in GGR and by RNAPII in TCR. After unwinding of the DNA around the lesion, the two pathways converge, leading to the incisions by the endonucleases XPF and XPG, whereafter the gap is refilled and the remaining nick religated. Refer to the text for more details. Figure adapted from (120).

1.3.8 ICL repair in the *Xenopus* system

To recapitulate ICL repair in a cell-free system, Räschle et al. first made use of *Xenopus laevis* egg extracts that support DNA replication of plasmids (84,125). To this end, High-Speed Supernatant (HSS) of egg cytoplasm is added to the DNA, which leads to the assembly of the pre-RC. The subsequent addition of nucleoplasmic extract (NPE) leads to a single round of DNA synthesis. In their studies, they used a plasmid containing a single

sequence-specific cisplatin ICL. Using this system, they were able to identify several steps of ICL repair. First, two DNA replication forks converge on the lesion, where their leading strands stall 20 to 40 nts away from it. Recently, it could be shown that when only one fork reaches the ICL, the replicative CMG helicase is not unloaded from the stalled fork and the repair is blocked, meaning that two forks have to converge at the site of ICL for the repair to be initiated in this system (126). In another publication, they reported BRCA1 to be important for this helicase-unloading step (41). After helicase unloading and a delay of about 20 min, one leading strand approaches to within one nt of the ICL. This step is performed by replicative polymerases (127). After another delay of about 30 min, this is followed by dual incisions to uncouple the DNA strands, lesion bypass, and extension of the nascent strand beyond the lesion. In their system, depletion of XPF, but not of MUS81 or FAN1 lead to a defect in ICL repair and they could further show that XPF-ERCC1 together with SLX4/FANCP was able to perform the unhooking step. Furthermore, the ubiquitinated FANCD2 promoted the recruitment of SLX4 and XPF-ERCC1 to the ICL (56). The lesion bypass was then performed by a complex of REV1 and Pol ζ, which was stimulated by the FA core complex but not by ID2, and induced a short mutagenic tract (127). The *Xenopus* egg extract system has proven to be a powerful model for the study of replication-dependent repair. It has to be mentioned, however, that the existence of only one origin of replication on a relatively small DNA molecule is not equivalent with the situation in the mammalian cell. Furthermore, it cannot be excluded that the pathways for ICL repair differ in frog and man.

1.3.9 Replication traverse of ICLs

To test whether the model described for the *Xenopus* system also holds true for mammalian genomes, Huang et al. developed a single-molecule technique to assess the encounter of replication forks at ICLs in living cells (128). They used photoactivated psoralen as an ICL-inducing agent, as it yields a high frequency of ICLs relative to monoadducts compared to other cross-linking drugs and because of the possibility to link psoralen to an antigen that can be later detected by IF (129). Their DNA fiber-based approach revealed that the major fraction of forks that stalled at the site of an ICL was consistent with replication on either side of the ICL linking the two parental strands. They could also show that the replication was continuous in over 50% of the cases, leading them to postulate a new way of replicating at an ICL being the traverse of such lesions. They also found about 20% of single-sided patterns (one fork stalled at the ICL) and a slightly lower amount of patterns that were in agreement with the dual fork collision model derived from the *Xenopus* egg extract system. This picture was the same in different cell lines tested and they could exclude the firing of dormant origins as a major contributor to this scheme. They calculated the time cost for replication on DNA with ICLs to be about 4.5 to 7 min depending on the cell line used. The time cost for replication over Angelicin (an equivalent of psoralen that can only form monoadducts) was only about 1 min. By using knock-out cell lines they could further show that the FANCM/MHF complex is important for the replication traverse of ICLs, where specifically their DNA binding

and translocase activities were needed. In contrast, neither the FA core complex, nor ubiquitinated FANCD2 were required for this process.

1.3.10 Replication-independent ICL repair

There is strong evidence that in mammalian cells ICL repair can happen in the absence of replication or recombination. First insights came from a site-specifically cross-linked reported plasmid that lacked an origin of replication. Therefore, ICL repair had to take place in a replication-independent manner (75). This pathway happens mainly in the G0/1 phase of the cell cycle and is dependent on NER proteins and TLS pols, particularly REV1 and Pol ζ (83,130,131). The repair of cisplatin ICLs was dependent on the TCR-NER specific protein CSB, suggesting that this is the preferred pathway for this type of damage, which is reflected also by the sensitivity of CSA and CSB cells to cisplatin (131,132). However, by looking at different types of ICL inducing agents, including MMC and psoralen/UVA, it becomes clear that probably both, TCR and GGR pathways play a role in the recognition and repair of all types of ICLs, but that some are preferably repaired by one pathway or the other depending on the structure (75,130,132-134). It is unclear what structural changes the GG-NER recognizes in the ICL. DDB1-DDB2, which is important for initial UV damage recognition and binds to these lesions, is not involved in the recognition of ICLs (130). An involvement of MMR proteins, such as MutSβ, in the replication-independent ICL repair pathway and specifically in the recognition step is also possible (135). ICL repair in G1 was independent of HR and RAD51, although some type of recombinational events occurred that was dependent on XPF in a cell-free cross-link repair synthesis assay (136). RPA and PCNA were also found to be required for the repair synthesis step as well as the TLS pols REV1 together with Pol ζ (131,137). A current model for mammalian G1 ICL repair suggests the recognition by GGR or TCR and a first round of unhooking the ICL by NER proteins, followed by TLS by REV1 and Pol ζ and finally another round of NER to remove the remnant of the ICL on the other strand (see **Figure 9**). It is unclear, however, whether the NER proteins fulfill the dual incisions to unhook the ICL alone or whether other proteins are also involved in this process (136,138,139).

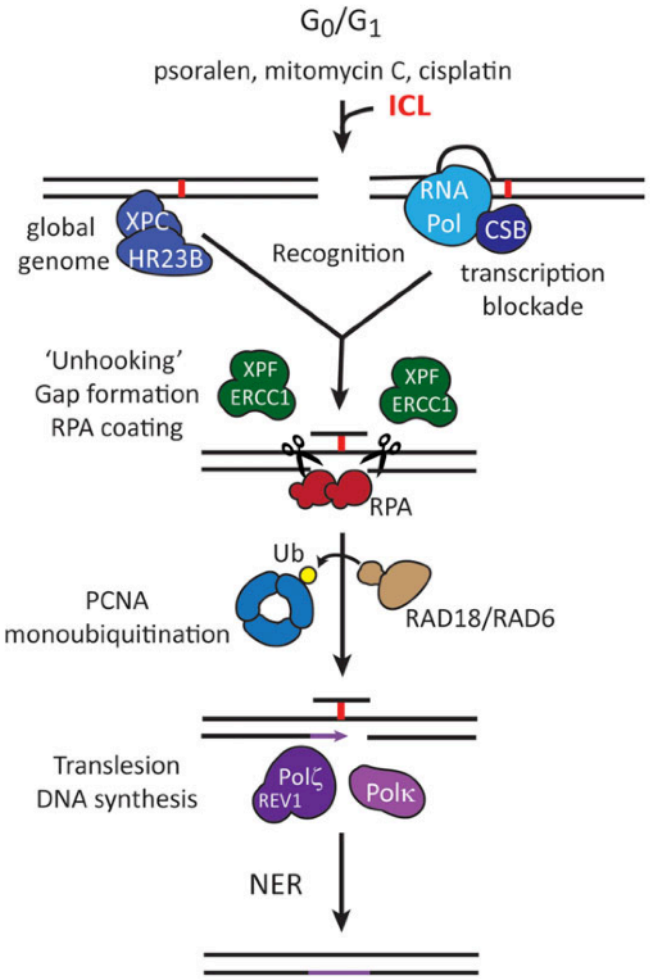

Figure 9: Replication-independent ICL repair in mammalian cells. Replication-independent ICL recognition in mammalian cells has some parallels with the NER pathway and can also be initiated in a global or transcription-coupled way. After unhooking, TLS pols refill the gap and a second round of NER removes the remaining adduct. Figure modified from (140).

1.3.11 ICL repair by glycosylases

Only recently, a novel DNA glycosylase from the bacterium *Streptomyces sahachiroi* was characterized, revealing some very interesting properties that could have tremendous impacts on the field of ICL repair (141). Azinomycin B is a novel ICL-inducing agent, leading to major groove ICLs at GNC or GNT sequences, isolated from *S. sahachiroi* and *Streptomyces griseofuscus*. Obviously, the producers have to protect themselves from the lethal effects of the drug. This is usually done by target site modifications, drug inactivation, drug binding or export (141). They identified Orf1 as an essential resistance protein against azinomycin B from the *S. sahachiroi* genome. Further characterization of the protein revealed that it binds to native DNA non sequence-specifically and interacts with an even greater affinity with azinomycin B-modified DNA. Further biochemical analysis showed that Orf1 could reverse the cross-links induced by azinomycin B, restoring the denaturation of the dsDNA under alkaline conditions. As no smaller DNA fragments were observed, the cleavage did not occur at the phosphodiester backbone. When the Orf1- dependent products were further incubated with Endonuclease IV (an apurinic/apyrimidinic (AP) endonuclease), this led to the degradation of

the DNA, indicating that Orf1 reverses the cross-link by depurination. Monoalkylated and interstrand cross-linked DNA molecules are both reversed by Orf1, which incised the adducted guanosines in both strands in the case of the ICL. If the two glycosidic incisions or the following AP endonuclease cleavages were temporally shifted, this pathway could even get around the formation of potentially toxic DSBs. This publication opened the possibility of a glycosylase-dependent ICL repair pathway, which might be similar to the BER pathway (see **Figure 10**). One problem with glycosylases regarding their usage in ICL repair is that they generally flip the damaged base out of the double helix and into their binding pocket for recognition and an ICL would be incompatible with base flipping. However, a recent publication shows the crystal structure of the first known DNA glycosylase that does not require base flipping for binding or excision, AlkB from *Bacillus cereus* (142). This further substantiates the probability of a glycosylase-dependent ICL repair pathway and it would be interesting to see whether Orf1 also works without base flipping, potentially by its crystal structure. It has to be seen whether an analogous pathway exists in mammalian cells, which is highly probable due to the simplicity of this type of pathway compared to, e.g. the FA pathway.

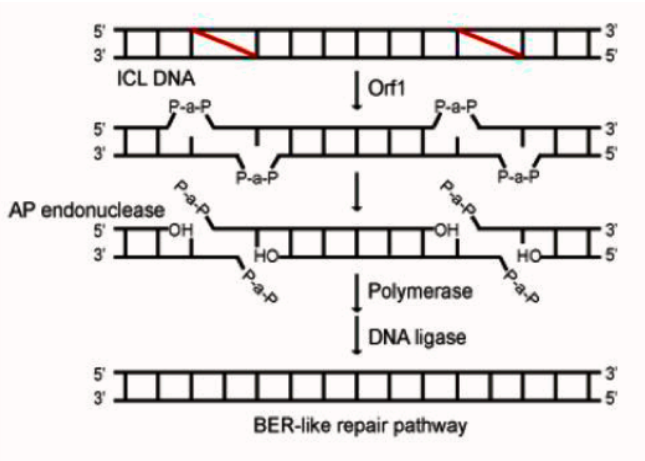

Figure 10: Potential BER-dependent ICL repair pathway identified in *S. sahachiroi*. Orf1, the newly identified glycosylase incises the glycosydic bond of the adducted guanines. An AP endonuclease cleaves at the resulting abasic site, leading to a nick that can be refilled by a polymerase before the DNA ligase restores the ds nature of the DNA. Figure taken from (141).

1.3.12 ICL repair in yeast

Most of the knowledge about ICL repair in yeast comes from genetic studies and, in accordance with the results from mammalian cells listed above, most of the mutants belong to one of the three pathways; NER, HR or TLS.

The first incision at the ICL is made by the NER enzymes Rad1-Rad10, Rad2, and Rad3 (homologs of human XPF-ERCC1, XPG and XPD). Rad1-Rad10 makes incisions 5' to the ICL, while Rad2 makes the incision 3' to it. ICL repair in yeast leads to a high amount of DSBs, which are then removed by HR. In yeast, all NER mutants are similarly sensitive to ICL agents, as opposed to mammalian cells, where XPF-ERCC1 seems to be exclusively important for the repair of ICLs (33).

As mentioned above, yeast have only one *SNM1* gene (also called *PSO2*), compared to the three homologs in human. *Pso2* deficient cells are exquisitely sensitive to ICL agents, suggesting that the protein does not have a role in another pathway. Pso2 has both, 5'->3' exonuclease and structure-specific endonuclease activities and is thought to act downstream of the initial incisions controlled by NER (63).

The only FANC core complex proteins that have homologs in *S. cerevisiae* are FANCM (Mph1) and FANCT. Other FANC proteins that have homologs in budding yeast are FANCJ (Chl1), FANCO (Rad51c), FANCP (Slx4), FANCQ/XPF (Rad1) and FANCR (Rad51) and the FANCM-associated factors MHF1 and MHF2 (Mhf1 and Mhf2). Interestingly, disruption of these FA-like genes in budding yeast did not sensitize them towards ICL agents, leading to the assumption that there is no FA or FA-like pathway in yeast. However, recently there have been two reports about the discovery of a FA-like pathway in budding yeast, which is S-phase dependent and plays a secondary role to the Pso2- and TLS-dependent pathway(s) (143-146).

In general, mammalian and yeast ICL repair are quite similar. The main differences are the involvement of FANC and BRCA proteins in mammals and the greater involvement of NER in yeast.

1.4 FEN1

As mentioned above in **chapter 1.2.1**, FEN1 (flap endonuclease 1) and DNA2 are the main nucleases responsible for flap removal during Okazaki fragment processing. In addition, FEN1 has a role in long-patch BER, where it cleaves flaps generated by Polβ's strand displacement activity. FEN1 is a metallonuclease, having a nuclease domain and an extended C-terminal region, which is important for interactions with PCNA and WRN (147,148). FEN1 preferably binds to a structure containing a 1-nt 3' flap and a 5' flap, which it recognizes at the base of the flaps (149). FEN1 then interacts with both dsDNA regions close to the flaps and bends the DNA at the site of the nick, where also the active site resides. Tight contacts with the 3' flap nucleotide by a hydrophobic wedge show the importance of this structure. The interactions with the 5' flap happen only at the base and DNA binding is then further stabilized by a threading mechanism, in which first the 5' end of the flap and then the entire flap passes through a gateway structure. Such a threading mechanism to select for flap structures is seen in other nucleases of the superfamily, such as XPG, GEN1 and EXO1. A cap structure, which is present only in EXO1 and FEN1, ensures that only ssDNA that contains an end passes through and leads to the selectivity (150). DNA2 recognizes ssDNA ends and uses a similar threading mechanism to directly displace RPA (151).

Additionally, it has been shown that FEN1 possesses a gap endonuclease (GEN) activity, which allows it to cleave stalled replication fork-like structures or bubbles opposite a gap or anick. The cleavage activity was dependent on its interaction with WRN (152). The threading mechanism seems to contradict such a GEN activity as no free end is present in these

structures, but differential regulation on several levels might ensure the proper usage of different FEN1 activities during various DNA metabolic processes (152).

In addition to its 5' flap and GEN activities, FEN1 possesses a 5'->3' exonuclease activity on dsDNA, for which it needs an entry point, but it can also degrade blunt ended DNA, albeit less efficiently (153). The exonuclease activity has been associated with apoptotic DNA fragmentation and Okazaki fragment processing (153).

FEN1 mutant MEFs (mouse embryonic fibroblasts) harboring a single mutation in the nuclease domain that abolishes the exonuclease and GEN but not the flap endonuclease activity (E160D) are sensitive towards the ICL inducing agent MMC (154). This could indicate that FEN1 is able to unhook or recover ICL-stalled replication forks and that this function is dependent on its GEN and not the flap endonuclease activity.

1.5 FAN1

1.5.1 FAN1 discovery

The first notion of FAN1 (FANCD2-associated nuclease 1) in our lab came from a MLH1 interactome study, where the two proteins were shown to strongly interact (155). FAN1 was then identified and studied in four different laboratories independently (156-160). The reasons for the interest in this protein were diverse; it appeared in a genome-wide shRNA screen as causing sensitivity to MMC (157) and MacKay et al. became interested in the protein due to its domain organization (159). All the studies reported the presence of a RAD18-like zinc finger, the UBZ domain at its N-terminus and a PD-D/E(X)K type endonuclease motif within a VRR-nuc (virus-type replication-repair nuclease) domain at the C-terminus. FAN1 is the only eukaryotic protein containing this VRR-nuc domain, which is mainly found in bacteriophages and bacteria as a standalone domain. It also possesses a putative DNA-binding (SAP (SAF-A/B, Acinus, and PIAS)) domain. There are orthologs of FAN1 in prokaryotes and many eukaryotes, however not in budding yeast. The four initial studies all reported the sensitivity of FAN1-depleted human cells or *C. elegans* strains towards ICL agents, like cisplatin, MMC and HN2, but not to other DNA damaging agents that were tested (camptothecin (CPT), UV, ionizing radiation (IR), methyl methanesulfonate (MMS), and APH) and an increase in chromosomal instability after treatment with those drugs. Furthermore, it was reported that FAN1 colocalizes with FANCD2 at ICL-induced subnuclear foci and that FAN1 acts downstream of the ID2 complex, as FANCD2- or FANCI-depleted cells showed a decrease in the formation of FAN1 foci. Moreover, FANCD2 containing a mutation at the site of monoubiquitination (K561R) was not able recruit FAN1 to repair foci, indicating the importance of this post-translational modification in the interaction between the two proteins. Consistently, FAN1's UBZ domain was necessary and sufficient for focus formation and an immunoprecipitation (IP) experiment showed the physical interaction between FAN1 and ID2. Moreover, Liu et al. showed the direct interaction between FAN1 UBZ domain and ubiquitin and that depletion of this domain leads to hypersensitivity of the cells to MMC. The direct

interaction of FAN1 with the ubiquitinated ID2 complex was not shown in these studies and the interaction of FAN1 with another ubiquitinated partner, activated through ID2 action still exists. There are differences in the results to whether FAN1 is also able to interact with the non-modified version of FANCD2.

In all four studies biochemical characterization of FAN1 was performed, which showed that it is a 5' flap endonuclease and a 5'->3' exonuclease. It has minor endonuclease activity on 3-way junctions and needs a gapped, nicked or recessed end as a starting point for its exonuclease activity, although it can also start at blunt ends with lower activity. It also consistently cleaves off the 5' label from the flap about 3-4 nts within the ssDNA.

Moreover, MacKay et al. and Kratz et al. showed that the formation of foci of DSB markers is normal in FAN1-depleted cells treated with MMC, but that the disappearance of those foci, in particular of RAD51, RPA and γH2AX, is slower. This would hint towards a function for FAN1 downstream of DSB formation, which was supported by the fact that depletion of FAN1 reduced the efficiency of I-SceI-dependent HR. However, sister chromatid exchange was not elevated upon FAN1 depletion (159).

1.5.2 Roles of FAN1 outside of ICL repair

As FAN1 was identified by means of its interaction with MMR proteins in our laboratory, the possibility of an importance for this interaction and therefore FAN1 involvement in MMR exists. MMR involves the degradation of the error-containing strand starting from a nick, an action that is mainly attributed to EXO1 (exonuclease 1). Indeed, there is evidence from our lab and others that FAN1 can replace EXO1 in the process of MMR *in vitro* and *in vivo* (161).

Lai et al. showed that FAN1 is degraded during mitotic exit, potentially by the anaphase promoting cyclosome complex (APC/C) via its activator Cdh1 (162). FAN1 might itself play a role in mitotic exit as this was delayed when overexpressing FAN1 in U2OS cells. Depletion by siRNA on the other hand, led to a faster progression through the cell cycle.

The first notion of FAN1 potentially having a role at stalled replication forks came from Shereda et al., who showed that FAN1 localizes to HU-induced foci to a similar degree as with MMC (163). Foci formation caused by both drugs was dependent on the UBZ domain and on FANCD2 monoubiquitination. In another more recent study, the authors show that FAN1 is involved in the recovery of stalled replication forks after APH treatment together with BLM and FANCD2 (164). This process was independent of FANCD2 monoubiquitination or FAN1 UBZ domain. The authors also observed a function of the complex in the suppression of new origin firing. Furthermore, they proposed another, FANCD2-dependent protective function in stalled fork stability by restricting inappropriate degradation by FAN1.

1.5.3 FAN1 in *Schizosaccharomyces pombe*

As mentioned before, yeast does not have the FA pathway for the repair of ICLs. Interestingly, FAN1 only has an ortholog in *Schizosaccharomyces pombe* (*S. pombe*) and not

in *Saccharomyces cerevisiae* (*S. cerevisiae*). This implies, as suggested by several previous studies, that FAN1 has a FA-independent role in ICL repair, which can be studied in yeast. Fontebasso et al. did genetic analyses to dissect the different pathways to resolve ICLs in *S. pombe* (165). First, they wanted to look for a potential interaction between MMR proteins and Fan1 in *S. pombe*, but they could not find any direct physical interactions and also the mutation rate was not increased in *fan1* deletion strains, arguing against a role of SpFan1 in MMR. Furthermore, *fan1* deletion led to sensitivity exclusively to ICL-inducing agents, which was markedly increased by co-deletion of *pso2*, suggesting that the two gene products act in parallel pathways during ICL repair in *S. pombe*. Rhp18, the fission yeast homolog of Rad18, and Fml1, the homolog of FANCM were epistatic to Fan1, while Rad13, the homolog of XPG was epistatic to Pso2. Rad16, the homolog of XPF on the other hand was epistatic to both, the Fan1- and Pso2- dependent pathways. HR was also needed in both pathways. This suggests that in fission yeast, there are two pathways to repair ICLs. One could involve the recognition of the damage by Fml1, followed by Rad16- and Fan1-dependent ICL processing and the subsequent re-fill of the DNA by TLS pols. A second pathway could rely on Rad16, Pso2 and Rad13 for lesion unhooking and both pathways would result in the induction of DSBs, which are repaired by HR. It has to be seen whether higher eukaryotes use similar pathways for the repair of ICLs.

1.5.4 FAN1 structure

Recently, there has been an unexpected and striking accumulation of publications regarding FAN1, specifically concerning its structure together with some biochemical characterization on cross-linked substrates (166-169). All structures were obtained with FAN1 variants lacking the UBZ domain. The common sense is that, as opposed to the prokaryotic stand-alone VRR-Nuc domain proteins, FAN1 does not support dimerization due to a helical insertion between two β sheets of the domain. Only one of the studies described dimerization of FAN1, which could also be an artifact of crystallization (169). The monomeric vs. dimeric forms helps to explain why the prokaryotic VRR-Nuc enzymes support HJ cleavage, whereas FAN1 prefers 5' flaps. Furthermore, FAN1 has four domains - an N-terminal domain (NTD), SAP, TPR (tetratricopeptide repeat) and VRR-Nuc domains. Each domain plays a specific role in pre-nick duplex, post-nick duplex, and 5' flap ssDNA recognition and in bending the DNA at the nick (see **Figure 11**). The 5' flap ssDNA is accomodated in a pocket at the interface of the NTD, TPR and VRR-Nuc domains. Next to the pocket, there is a hole between the TPR and VRR-Nuc domains, which would be big enough to accomodate ssDNA and not dsDNA. As FEN1 and EXO1 contain a hole where the 5' flap can enter, the same could be expected for FAN1. However, the flap does not pass through this hole in FAN1 but rather makes a turn in the opposite direction, which is wide open (166).

Gwon et al. showed that the two-metal active site is positioned 5 nt downstream from the pocket that accommodates the 5' flap, explaining the flap cleavage site by FAN1 being a few nts within the post-nick duplex (166). The cleavage site was dependent on the position of the

junction between the pre-nick duplex and the ssDNA and not on the flap or any extension of the 3' nicked strand.

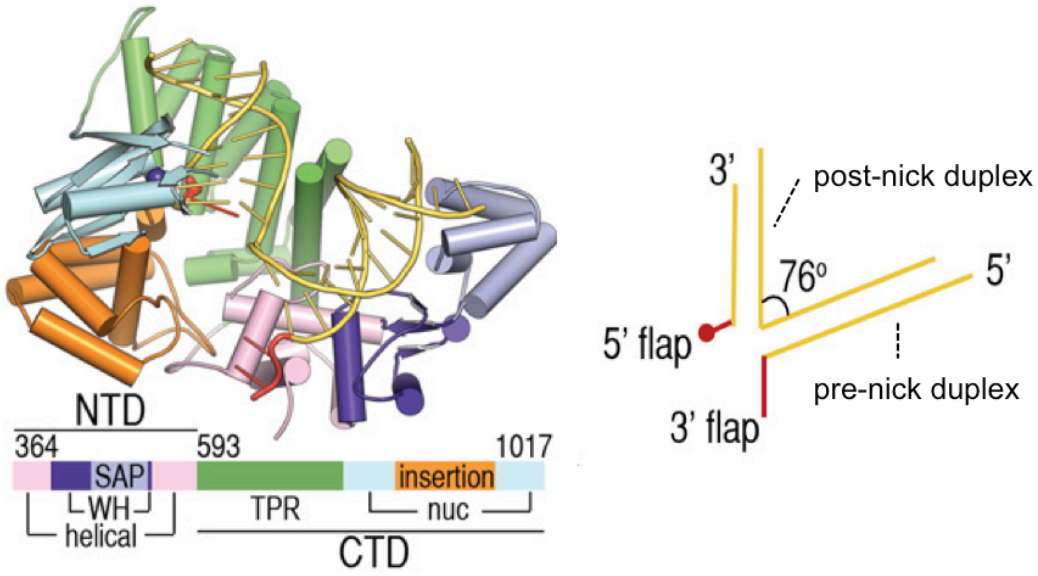

Figure 11: FAN1 structure. The different domains of FAN1 are colored according to the graphical representation below. A line drawing of the DNA molecule is shown on the right with the same color code as in the cartoon of the structure. Figure adapted from (168).

Wang et al. performed nuclease assays under physiological salt conditions and showed that the addition of a 3' flap increased the FAN1 binding affinity and this was explained by close contacts between the first two nts from the 3' flap with the protein seen in the crystal structure (168). Under their conditions there was a much decreased affinity and nuclease activity of FAN1 when incubated with 5' flaps longer than two nts compared to the shorter ones. They could further show that the 'exonuclease' activity is achieved by sequential (endonucleolytic) cuts resulting in 3 nt products. However, the sequential incisions stalled after four cuts, suggesting that a gap length beyond 12 nt is no longer compatible with FAN1 binding. They suggest that, in case of a complementary 3' flap or a polymerase in the cell, this could close the emergent gap and restore FAN1 binding and processivity.

Two of the studies showed that a stable ICL (one consisting of two cytosines linked through a triazole moiety at their N4 positions and the other being a psoralen ICL) could be bypassed by FAN1, when the cross-link was located 6 nt away from the nick within the post-nick duplex. It cleaved both immediately 5' to the ICL and immediately 3' to it, and cleavage was less efficient when the cross-link was moved further away from the junction. This suggests that FAN1's active site is relatively flexible and can deal with possible distortions caused by the ICL in the DNA (168,169).

1.5.5 FAN1 in diseases

As FAN1 depletion leads to hypersensitivity to cross-linking agents and it interacts with FANCD2, it was tempting to speculate that also FAN1 is encoded by a FA gene. Several groups addressed this question, e.g. Trujillo et al., who analyzed 4 patients with homozygous deletions of a region on chromosome 15, where FAN1 is located (170). The cells derived from the patients (where the lack of FAN1 expression was confirmed), scored negative for the classical tests for FA and also the symptoms of the patients were different from FA, which included microsomy, microcephaly and neurodevelopmental abnormalities, but no anemia, bone marrow failure or skeletal abnormalities were reported. These data, together with the notion that FAN1 was not epistatic to the FA pathway in chicken cells (156), exclude *FAN1* as being a FA gene and suggests that the protein processes only a specific subset of ICLs, which are not addressed by the classical FA pathway. This would imply the existence of other nuclease(s) having redundant roles with FAN1.

Zhou et al. instead reported that mutations in FAN1 are the cause of KIN (karyomegalic interstitial nephritis) in several affected families (171). 9 out of 10 families with KIN analyzed by sequencing had mutations in both alleles of FAN1. Also here, they report the negative score of the patient-derived cell lines in the typical FA test and the mutant that is defective in the interaction with FANCD2 (C44A/C47A) fully rescued the MMC sensitivity in the FAN1 deficient cells, in contrast to the nuclease-dead mutant (D975A/K977A). They also depleted FANCD2, SLX4, XPF or MUS81 from FAN1 mutated fibroblasts and showed that the MMC sensitivity was greater than the single depletions, which suggests that FAN1 acts outside the FA pathway in a redundant fashion. Protein levels of FAN1 are particularly high in the kidney, which might explain the differential phenotypes by the dependencies of different organs on FAN1 or the FA pathway to repair ICL damage.

FAN1 variants were further associated with schizophrenia and autism (172) as well as with MMR-proficient hereditary nonpolyposis colorectal cancer (173). Smith et al. just recently suggested FAN1 as a candidate susceptibility gene for high-risk pancreatic cancer through an exome sequencing study performed on these cancers (174).

1.6 My aim

Depletion of FAN1 renders cells sensitive towards ICL agents. Given its endo- and exonuclease activites, a possible function for FAN1 lies in the unhooking step, where endonucleases incise the DNA 5' and 3' of the ICL on the same strand. We envisioned that FAN1 could itself make both cleavages at either side of the cross-link, by first using its endo- and then its exonuclease activity to bypass the lesion. To test this hypothesis, I used synthetic ICL-containing DNA substrates that were generated in the laboratory of Prof. Orlando D. Schärer. The ICL resembles a NM-derived ICL, although its structure is less DNA distorting. By incubating purified FAN1 with these substrates, I could test if and how FAN1 processes DNA molecules bearing such an ICL. Furthermore, I made use of cell biological assays to

study the *in vivo* role of FAN1 in ICL repair. Pulse-field gel electrophoresis allowed me to assess the amount of DSBs in the cells after treatment with cross-linking agents. This was used as a surrogate for ICL unhooking as the uncoupling of ICLs during S-phase leads to the formation of DSBs and cells proficient and deficient for FAN1 were compared. In addition, I analyzed the functional redundancy between FAN1 and MUS81 by survival studies. The results from these studies were published in 2015 in the Journal of Biological Chemistry and are listed in section 3.1. Moreover, I performed studies on potential redundancies between the two endonucleases FAN1 and FEN1 in ICL repair. FEN1 shows a much higher activity in extracts on model DNA 5' flap structures than FAN1 and there is evidence that FEN1 could have a role in ICL repair. The results to this part are listed in section 3.2.

2 Results

2.1 FANCD2-associated Nuclease 1, but not Exonuclease 1 or Flap Endonuclease 1, is able to unhook DNA interstrand cross-links *in vitro*

Article published in 'the Journal of Biological Chemistry', 2015

Authors:

Julia Pizzolato, Shivam Mukherjee, Orlando D. Schärer, and Josef Jiricny

I contributed to this work by performing all the experiments included, helping to design them and co-writing the manuscript.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 37, pp. 22602–22611, September 11, 2015

FANCD2-associated Nuclease 1, but Not Exonuclease 1 or Flap Endonuclease 1, Is Able to Unhook DNA Interstrand Cross-links *in Vitro*[*]

Received for publication, May 6, 2015, and in revised form, July 27, 2015 Published, JBC Papers in Press, July 28, 2015, DOI 10.1074/jbc.M115.663666

Julia Pizzolato[‡], Shivam Mukherjee[§], Orlando D. Schärer[§¶], and Josef Jiricny[‡‖1]

From the [‡]Institute of Molecular Cancer Research, University of Zurich and the [‖]Department of Biology, Swiss Institute of Technology (ETH) Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland, and the Departments of [§]Chemistry and [¶]Pharmacological Sciences, Stony Brook University, Stony Brook, New York 11794-3400

Background: We studied how the endo/exonucleases EXO1, FAN1, and FEN1 process substrates resembling replication forks blocked by interstrand cross-links (ICLs).
Results: All three enzymes cleaved off the single-stranded 5′ flap, but FAN1 was also able to incise the substrate behind the ICL.
Conclusion: FAN1 can unhook ICLs.
Significance: *In vivo*, FAN1 may not require a 3′ flap nuclease to unhook ICLs.

Cisplatin and its derivatives, nitrogen mustards and mitomycin C, are used widely in cancer chemotherapy. Their efficacy is linked primarily to their ability to generate DNA interstrand cross-links (ICLs), which effectively block the progression of transcription and replication machineries. Release of this block, referred to as unhooking, has been postulated to require endonucleases that incise one strand of the duplex on either side of the ICL. Here we investigated how the 5′ flap nucleases FANCD2-associated nuclease 1 (FAN1), exonuclease 1 (EXO1), and flap endonuclease 1 (FEN1) process a substrate reminiscent of a replication fork arrested at an ICL. We now show that EXO1 and FEN1 cleaved the substrate at the boundary between the single-stranded 5′ flap and the duplex, whereas FAN1 incised it three to four nucleotides in the double-stranded region. This affected the outcome of processing of a substrate containing a nitrogen mustard-like ICL two nucleotides in the duplex region because FAN1, unlike EXO1 and FEN1, incised the substrate predominantly beyond the ICL and, therefore, failed to release the 5′ flap. We also show that FAN1 was able to degrade a linear ICL substrate. This ability of FAN1 to traverse ICLs in DNA could help to elucidate its biological function, which is currently unknown.

DNA interstrand cross-links (ICLs)[2] are highly cytotoxic because of their propensity to block both transcription and replication (1). In all organisms studied to date, ICL repair involves protein constituents of the nucleotide excision repair pathway, translesion polymerases, and the machinery of homologous recombination (2). In higher eukaryotes, ICL processing also requires proteins of the FANC complementation group and their interactors. Correspondingly, cells lacking these proteins are exquisitely sensitive to ICL-inducing agents. This is of substantial clinical relevance. Mutations in genes encoding FANC polypeptides predispose to Fanconi anemia (FA), a severe genetic disorder characterized by bone marrow failure, congenital abnormalities and cancer predisposition, and cellular sensitivity to the cross-linking agents diepoxybutane or mitomycin C (MMC) is used in FA diagnosis (3).

Of the 17 FANC proteins identified to date, eight form the so-called core complex, which is activated at stalled replication forks by the ATR kinase. The FANCL subunit of the activated core complex monoubiquitylates the FANCD2/FANCI heterodimer that is believed to recruit to the cross-linked site in chromatin the polypeptides required for ICL removal (4). In the first step, these proteins have to separate the cross-linked strands in a process referred to as "unhooking," in which one strand of the duplex is nucleolytically incised on either side of the cross-link. Studies carried out during the past three decades helped to identify several nuclease-encoding genes, the disruption of which results in sensitivity to ICL-inducing agents, and biochemical characterization of their respective gene products confirmed that these nucleases could indeed be involved in ICL unhooking. The primary candidates are SLX1-SLX4, XPF-ERCC1, MUS81-EME1, SNM1A, and FAN1. Despite some seminal mechanistic studies that have been described in the recent literature (reviewed in Ref. 5), our understanding of the possible functions of these enzymes in ICL unhooking is still rudimentary. The situation is further complicated by functional redundancies of several of the involved proteins, their interplay with other members of the ICL repair pathway, and differences in helical distortions induced by different types of ICLs (6).

FANCD2-associated nuclease 1 (FAN1) has been identified independently in four laboratories (7–10). It possesses a conserved PDX_n(D/E)XK active site motif common to the restric-

* This work was supported by Swiss National Science Foundation Grants 310030B-133123 and 31003A-149989 and by Advanced ERC Grant MIRIAM (to J. J.). Funding for open access charge was from the University of Zurich. The authors declare that they have no conflicts of interest with the contents of this article.
[1] To whom correspondence should be addressed: Institute of Molecular Cancer Research of the University of Zurich and Department of Biology of the Swiss Institute of Technology (ETH) Zurich, Winterthurerstr. 190, 8057 Zurich, Switzerland. Tel.: 41-44-635-3450; Fax: 41-44-635-3484; E-mail: jiricny@imcr.uzh.ch.
[2] The abbreviations used are: ICL, interstrand cross-link; FA, Fanconi anemia; MMC, mitomycin C; ss, single-stranded; ds, double-stranded; NM, nitrogen mustard; DSB, double strand break.

tion nuclease-like superfamily. This site is embedded in a VRR-nuclease domain, the structure of which has been solved recently (11). FAN1 has been implicated in ICL repair because its knockdown (7–10) or knockout (12, 13) caused sensitivity to the ICL-inducing agents MMC and cisplatin. The protein contains also a ubiquitin-binding zinc finger domain, which is necessary and sufficient for its recruitment to MMC-induced foci of the monoubiquitylated form of FANCD2/I (7–10). As a consequence, FAN1 was predicted to be a novel FANC protein. However, experiments with chicken DT40 knockout cell lines failed to yield evidence of an epistasis between FAN1 and FANCC or FANCJ (13), and no mutations in *FAN1* were identified in unassigned FA patients. Instead, FAN1 loss has been linked to karyomegalic interstitial nephritis (14). Therefore, even though its involvement in ICL processing is beyond doubt, FAN1 does not appear to be a FANC protein. However, current evidence does not exclude the possibility that FAN1 does participate in the FANC pathway but that its deficiency does not lead to FA because of functional redundancy with other polypeptide(s).

In an attempt to address the latter point, we set out to compare how structures resembling ICL-blocked replication forks are addressed by three 5′ flap endonucleases and 5′ to 3′ exonucleases with known roles in DNA replication and repair: EXO1, FAN1, and FEN1. EXO1 (15) and FEN1 (16) belong to the same enzyme family and have been reported to incise 5′ flap structures primarily one nucleotide 3′ from the single-strand (ss)/double-strand (ds) junction (for a review, see Ref. 17). FAN1 has been shown to incise 5′ flap structures two to four nucleotides 3′ from the ss/ds junction (9). Given their similar substrate specificities, we wondered whether these enzymes could compensate for one another in ICL processing. We therefore synthesized a 5′ flap structure containing a single cross-link resembling an ICL induced by nitrogen mustards (NMs), positioned two nucleotides 3′ from the ss/ds junction (18). As controls, we also synthesized a linear duplex containing a similar ICL as well as several unmodified substrates. We now show that all three enzymes could cleave the 5′ flaps but that FAN1 was also able to traverse the cross-link on both ICL substrates. Furthermore, we show that FAN1 depletion leads to a reduction in the number of MMC-induced double-strand breaks (DSBs). This could imply that FAN1-dependent processing of ICLs leads to the generation of repairable DSBs, whereas FAN1 absence or malfunction may lead to ICL persistence and cytotoxicity.

Experimental Procedures

Antibodies—Sheep α-KIAA1018/MTMR15 (a gift from J. Rouse, 1:625), rabbit α-TFIIH (Ser-19, Santa Cruz Biotechnology, 1:1000), mouse α-MUS81 (Sigma, 1:2000), mouse α-Lamin B1 (Abcam, 1:100), rabbit α-γH2AX (Ser-139, Cell Signaling Technology, 1:1000), mouse α-RPA2 (Calbiochem, 1:50), mouse α-CtIP (Asp-4, Santa Cruz Biotechnology, 1:250), and rabbit α-pRPA2 (Ser-4/Ser-8, Bethyl Laboratories, 1:500) were used.

siRNA Treatments—The following siRNAs (Microsynth, Balgach, Switzerland) were used: siFAN1, GUAAGGCUCUUUCAACGUA; siLuciferase, CGUACGCGGAAUACUUCGA; and siMUS81, CACGCGCTTCGTATTTCAGAA and CGGGAGCACCTGAATCCTAAT. The siRNA treatments were carried out 1 day after seeding the cells at ~30% confluency in 10-cm dishes. The transfection agent was Lipofectamine RNAiMAX™ (Invitrogen), which was used according to the instructions of the manufacturer.

Synthesis of Substrates—DNA ICLs were synthesized as described in Refs. 19, 20. Unmodified oligonucleotides were purchased from Microsynth. Sequences and graphical representation of the substrates can be found in Figs. 1*A*, 2*A*, 3*A*, and 4*A*.

Labeling and Annealing of Unmodified Substrates—The 5′ labeled flap substrate shown in Fig. 1*A* was generated by annealing oligonucleotide I labeled at its 5′ terminus with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (Hartmann Analytic) to the unlabeled oligonucleotides II and III at 95 °C and cooling down to room temperature in annealing buffer (25 mM Hepes-KOH (pH 7.4) and 50 mM KCl).

The 5′ labeled flap substrates shown in Fig. 4*A* were generated by annealing either the 5′-labeled oligonucleotide VII with the unlabeled oligonucleotides IX, XI, and XII or the 5′-labeled oligonucleotide VIII with the unlabeled oligonucleotides X, XI, and XII. The 3′-labeled substrates were generated by first annealing oligonucleotides I and IV (Fig. 2*A*) or V and VI (Fig. 3*A*), and filling in the overhangs using Klenow polymerase (New England Biolabs) and [α-^{32}P]dTTP (Hartmann Analytic) together with dATP, dCTP, and dGTP, followed where required by reannealing with oligonucleotide III (Fig. 2*A*).

Labeling and Annealing of ICL Substrates—The 5′ flap substrates were generated by labeling the 5′ termini of the cross-linked Y structure (oligonucleotides I and II, Fig. 1*A*) with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP and annealing with oligonucleotide III as described above. The 3′-labeled cross-linked substrates (Figs. 2*A* and 3*A*) were generated by filling in the overhangs using Klenow polymerase (New England Biolabs) and [α-^{32}P]dTTP together with dATP, dCTP, and dGTP, followed by annealing with oligonucleotide III (Fig. 2*A*) as described above. The recessed substrate (5′ CCCTCTTCTXTCCTTCTTTC 3′/5′GAAAGAAGXACAGAAGAGGGTACCATCATAGAGTCAGTG 3′, where X represents the cross-linked guanines, Fig. 3*F*) was labeled at the ds 3′ end by cordycepin (PerkinElmer Life Sciences) and terminal transferase (Roche).

In Vitro Nuclease Assays—WT and mutant FAN1 were purified as described previously (9) and dialyzed against nuclease buffer (25 mM Hepes-KOH (pH 7.4, FAN1, or pH 7.8, EXO1 and FEN1), 25 mM KCl, 1 mM $MgCl_2$, 0.5% PEG, and 0.05 mg/ml BSA) for 15 min at 4 °C before use. WT and mutant recombinant purified human EXO1 was a gift from Stephanie Bregenhorn (21). FEN1 was purchased from Gentaur (catalog no. C140). Unless stated otherwise, the nuclease assays were carried out by incubating the substrates for 30 min at 37 °C in the appropriate nuclease buffer (see above). Aliquots were withdrawn at the indicated time points, mixed with an equal volume of loading dye (80% formamide, 50 mM Tris-HCl (pH 8.3), and 1 mM EDTA), and loaded on 20% denaturing polyacrylamide gels (20 × 20 × 0.05 cm) that were run in TBE buffer for 1 h at 800 V. The gels were fixed, dried, and exposed to phospho screens,

which were then scanned in a PhosphorImager (Typhoon 9400, GE Healthcare). Quantification of the bands was carried out using ImageJ, and the graphs were generated using GraphPad Prism. The overall efficiency of substrate processing was obtained by dividing the integral values of all the product bands by the integral value of all the bands. The cleavage efficiencies of the 5′ flaps in the unmodified and ICL substrates were calculated by dividing the integral values of the product bands by the integral value of all bands.

Pulsed Field Gel Electrophoresis—Pulsed field gel electrophoresis was performed as described previously (22) with minor modifications. 24 h after MMC treatment (3 μg/ml), cells were harvested, and agarose plugs containing 250,000 cells/plug were generated. Quantifications were done using ImageJ, and graphs were produced using GraphPad Prism.

5-Ethynyl-2′-deoxyuridine (EdU) Staining and FACS Analysis—After treating the cells with 10 μM EdU (Life Technologies) for 30 min, they were harvested, fixed in 1% formaldehyde, and labeled using the Click-iT® cell reaction buffer kit (Life Technologies) according to the instructions of the manufacturer. DNA was stained with 1 μg/ml DAPI, and the samples were analyzed on a Cyan ADP flow cytometer (Beckman Coulter) fitted with Summit software v4.3 (Beckman Coulter).

Clonogenic Survival Assay—Survival assays were performed as described previously (9). Eight days after treating the cells with MMC or carmustine (bromo-chloro-nitrosourea, BCNU) at the indicated dosages, the cells were fixed and stained with 0.5% crystal violet in 20% ethanol, and colonies containing more than 50 cells were counted. MMC-containing medium was replaced with fresh medium after 24 h.

Results

An NM-like ICL Does Not Block the Endonuclease Activity of FAN1 on a 5′ Flap Substrate—The ability of FAN1 to cleave substrates containing 5′ flaps, coupled to the sensitivity of FAN1-depleted cells to ICL-inducing agents (7–10), indicates that this protein might be involved in ICL unhooking. Experimental evidence implicating FAN1 in this process has, however, not been forthcoming. In *Xenopus laevis* egg extracts, FAN1 depletion caused no detectable defect in unhooking of a single ICL in a plasmid substrate, but this could be explained by functional redundancy with other nucleases in the extract (23). To learn with which nuclease FAN1 could be redundant, we set out to test its ability to process a 5′ flap substrate containing a single NM-like ICL (Fig. 1*A*) and compared it with those of two other 5′ flap endonucleases, EXO1 and FEN1. FAN1-catalyzed cleavage of the control, unmodified 5′ flap substrate labeled at the 5′ terminus of the flap gave rise to several species between 11 and 21 nucleotides in length, the most prominent of which was a 19-mer arising as a product of endonucleolytic cleavage of the lagging (flap-containing) strand four nucleotides 3′ from the ss/ds junction (Fig. 1*B*, *lanes 1–6*, products **a**).

EXO1 and FEN1 belong to the same nuclease family, possess highly conserved active sites, and have been reported to cleave 5′ flaps preferentially one nucleotide 3′ from the ss/ds junction (17). They were therefore both expected to release 16-mer flaps from the unmodified substrate (Fig. 1*A*). Somewhat surprisingly, the flap fragments generated by the two enzymes differed. The main product of EXO1-catalyzed cleavage was a 17-mer, whereas FEN1 generated predominantly a mixture of 15-mer and 16-mer fragments (*cf.* Fig. 1, *C* and *D*, *lanes 1–6*). Both enzymes are known to bind to the template strand and bend the DNA at the break. Our result suggests that the 90° bend introduced by EXO1 binding (17) partially melted the flap-containing strand to generate a single nucleotide gap between the 3′ terminus of the leading strand primer and the ss/ds boundary and subsequently cleaved the substrate 3′ from the first nucleotide of the duplex. The FEN1-generated products can also be explained by its *modus operandi*. It bends the substrate DNA by 100°, melts two base pairs downstream of the ss/ds boundary, and cleaves the strand between these two extruded nucleotides. This would explain the origin of the 16-mer fragment. The origin of the 15-mer can also be explained. FEN1 prefers to generate directly ligatable products and achieves this by binding and bending the DNA so that the 3′ end of the leading strand primer melts to give rise to a single-nucleotide 3′ flap. The ss/ds boundary in this structure is shifted by one nucleotide in the 5′ direction. Cleavage of this 5′ flap between the two extruded nucleotides would then liberate a 15-mer flap and leave behind a nick that can be ligated.

Analysis of products generated by the three nucleases from the 5′ phosphorylated cross-linked substrate was slightly complicated by the fact that the oligonucleotide had to be labeled by polynucleotide kinase after ICL synthesis, which means that both 5′ termini were labeled even though the ds end was phosphorylated less efficiently than the ss flap. This increased the number of possible products shown in Fig. 1*A*, *right panel*. FEN1 flap endonuclease activity was only weakly inhibited by the presence of the ICL (Fig. 1*D*, *bottom panel*). Because this enzyme lacks exonuclease activity on blunt-ended, double-stranded DNA, product **b'** was clearly detectable, as were the released flap products (**a'**), of which the 15-mer predominated (Fig. 1*D*, *lanes 7–12*). EXO1 activity on the ICL substrate was inhibited (Fig. 1*C*, *bottom panel*), as judged by the lower intensity of bands **a'** compared with bands **a**. Interestingly, EXO1 failed to release the 17-mer flap from the ICL substrate (Fig. 1*C*, *lanes 7–12*).

As shown in Fig. 1*B*, *lanes 7–12*, the abundance of flap fragments **a**/**a'** arising through cleavages 5′ from the ICL by FAN1 was almost comparable with the unmodified and cross-linked substrates, as shown by the intensity of the 15–17-mer bands (*red arrowheads*) relative to the full-length substrate (Fig. 1*B*, *bottom panel*). Because of the fact that the lower substrate strand was also labeled, products **d'** (Fig. 1*A*) were also detectable in the autoradiograph (Fig. 1*B*, *lanes 7–12*). Products **c'** were generated by a 5′-to-3′ degradation from the blunt-ended 5′ terminus of the ICL substrate, but the nature of band **d'** was uncertain. However, given that it migrated only slightly slower than the 37-mer, we postulated that it might correspond to species **d'**, shown in Fig. 1*A* (see also below).

The most notable difference between the FAN1-generated products of the unmodified and the ICL substrate, however, was the absence of the 19-mer product in the latter reactions (Fig. 1*B*, *cf. lanes 1–6* and *7–12*). On the basis of the cleavage pattern of the unmodified flap substrate, this incision would

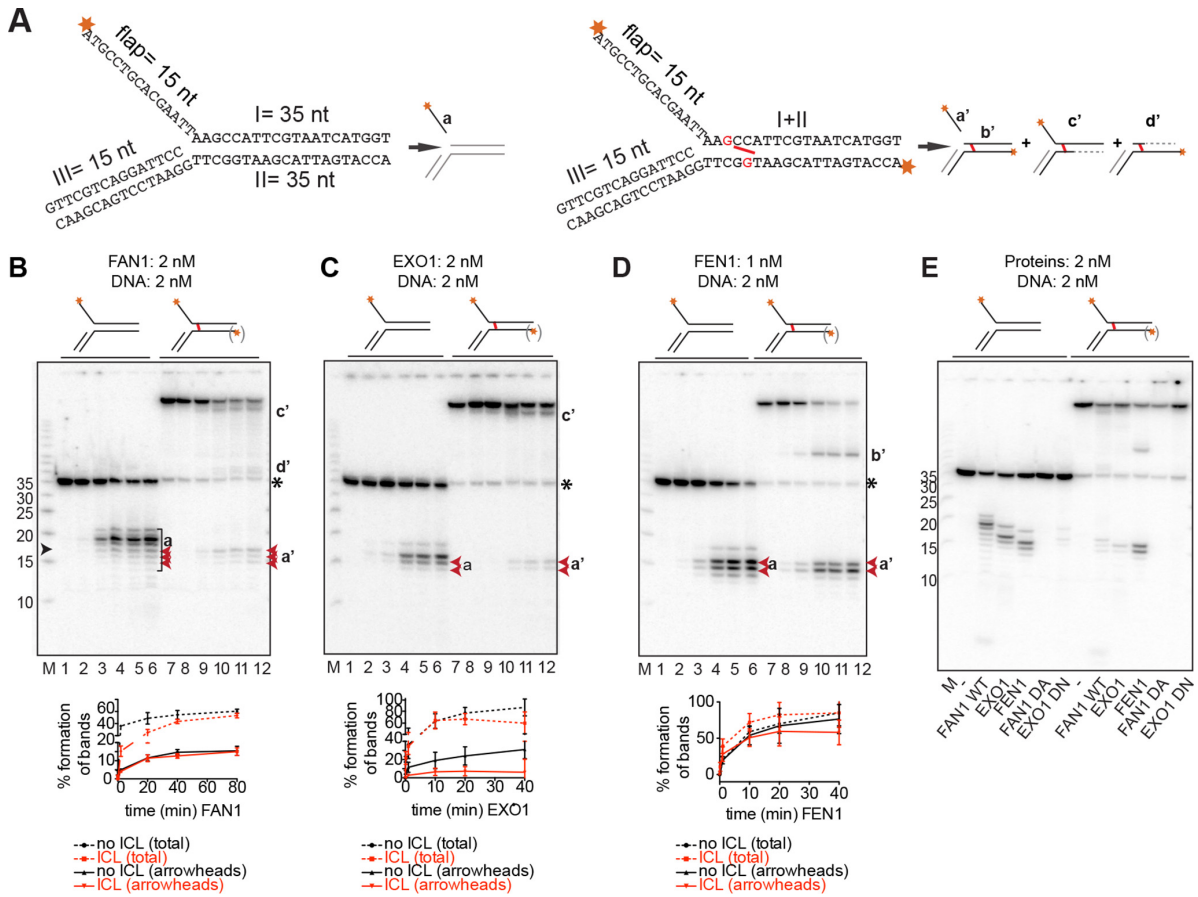

FIGURE 1. **Comparison of 5′ flap endonuclease activities and the specificities of FAN1, EXO1, and FEN1.** *A*, schematic representation of the DNA substrates used in this study. The unmodified flap substrates (*left panel*) were generated as described under "Experimental Procedures," and the labeling of the flap substrates with a nitrogen mustard-like interstrand cross-link (*right panel*, cross-link shown in *red*) resulted primarily, but not exclusively, in the labeling of the I strand. The likely products generated by the above enzymes from these substrates are also shown. *Red asterisks* indicate the positions of the ^{32}P-labeled 5′ phosphates. Fragments invisible on the autoradiograph are shown in *gray*. *Dashed lines* represent regions of exonucleolytic degradation. *nt*, nucleotides. *B–D*, product generated upon incubation of the substrates shown above the panels with FAN1 (*B*), EXO1 (*C*), and FEN1 (*D*). Aliquots were withdrawn at 0, 10 s, 2 min, 20 min, 40 min, and 80 min (*B*, *lanes 1–6* and *7–12*, respectively) or at 0, 10 s, 1 min, 10 min, 20 min, and 40 min (*C* and *D*, *lanes 1–6* and *7–12*, respectively). The protein-to-DNA ratios were 1:1 for FAN1 and EXO1 and 1:2 for FEN1. *M*, low molecular weight marker (Affymetrix). The oligonucleotide sizes are indicated on the *left*, and the position of the cross-link is indicated by a *black arrowhead*. The *lowercase letters* on the *right* correspond to the products in *A*. (Products **b′–d′** are seen solely in the reactions using the cross-linked substrate, in which both strands were labeled.) The *black asterisk* indicates the position of the non-cross-linked 35-mer oligonucleotide that was present in small amounts in the ICL substrate. The graphs below the autoradiographs of 20% denaturing polyacrylamide gels represent the quantification of either all product bands (*total*) or of only the bands indicated in *red* (*arrowheads*). The most prominent product bands in the reaction of FAN1 with the unmodified substrate resulted from incisions beyond the position of the ICL and are therefore not detectable in digestions of the cross-linked substrate. *Error bars* show mean ± S.D. ($n = 3$). *E*, comparative analysis of digestions of the indicated 5′ flap substrates by the three structure-specific endonucleases, including the nuclease-dead mutants of FAN1 (D960A) and EXO1 (D173N). DNA and proteins were all in equimolar ratios, and the reactions were carried out for 30 min at 37 °C.

have occurred on the 3′ side of the cross-link, and, therefore, the labeled flap fragment would not have been released from the substrate. In the experiments described above, we used recombinant FAN1 and EXO1 expressed in Sf9 cells and bacteria, respectively. To ensure that the observed enzymatic activities were indeed due to these polypeptides rather than contaminating nucleases, we also generated nuclease-dead variants of FAN1 (D960A) and EXO1 (D173N). As shown in Fig. 1*E*, the inactive variants generated no products comparable with those generated by the wild-type proteins.

To verify the incision on the 3′ side of the cross-link experimentally, we labeled the substrates at the 3′ termini of the upper strand (Fig. 2*A*). As anticipated, FAN1 (7–10), EXO1 (16), and FEN1 (24) degraded the unmodified substrate exonucleolytically from the sites of their respective endonucleolytic incisions in a 5′-to-3′ direction, with EXO1 (Fig. 2*C*, *lanes 1–6*) being more efficient than FAN1 (Fig. 2*B*, *lanes 1–6*), which was, in turn, more efficient than FEN1 (Fig. 2*D*, *lanes 1–6*). Most of the products detected in the FAN1-catalyzed degradation were longer than 11 nucleotides, which suggested that this enzyme might require a dsDNA stretch of ~10 base pairs to which to bind. In contrast, EXO1 appeared to completely degrade the incised strand. FEN1 predominantly generated product **a** from the unmodified substrate. Incubation of the ICL substrate with EXO1 generated species **c'** as the primary product, and lower amounts of species **a'** and **d'** (Fig. 2*C*, *lanes 7–12*), which suggests that this enzyme may prefer a blunt-end terminus to a 5′ flap as a substrate (15). FEN1 predominantly yielded product **a'**

FAN1-catalyzed Unhooking of ICLs

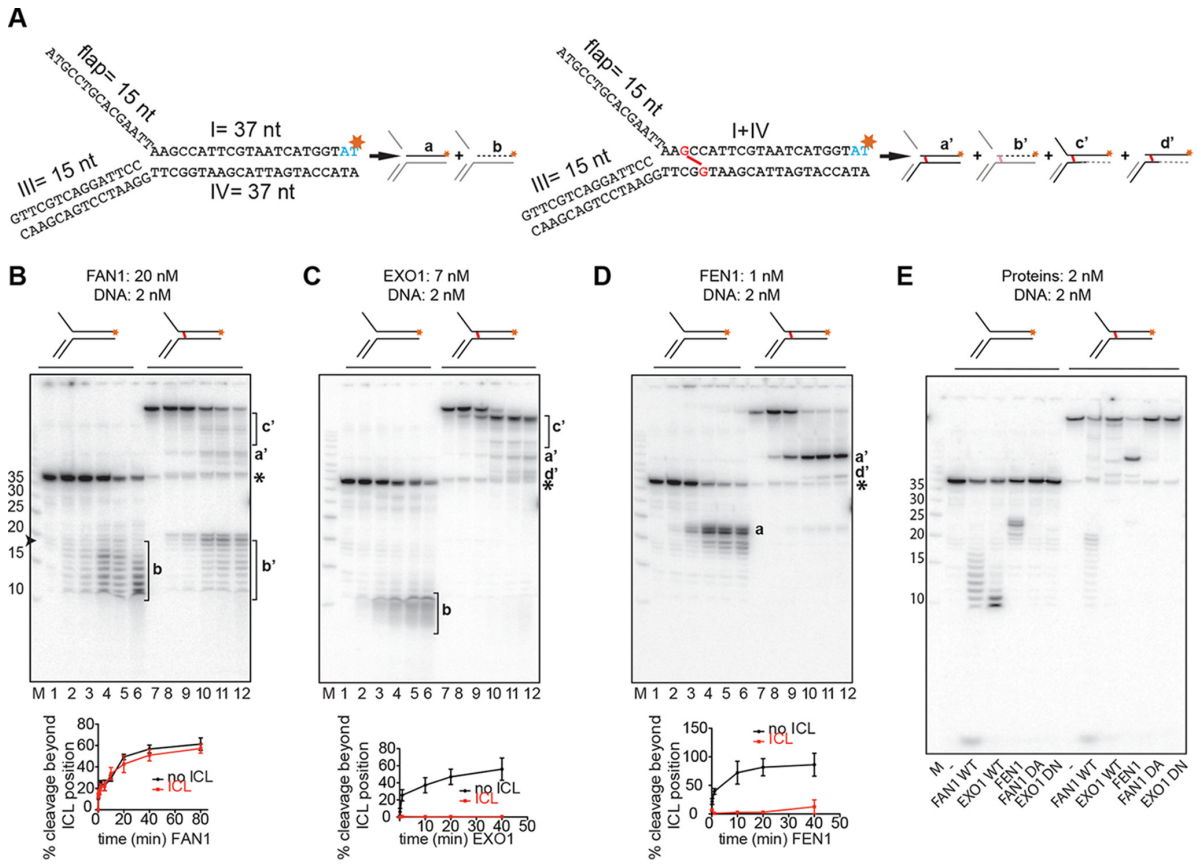

FIGURE 2. **Comparison of 5′ to 3′ exonuclease activities and the specificities of FAN1, EXO1, and FEN1.** *A*, schematic of the DNA substrates used in this study. The unmodified flap substrates (*left panel*) and the substrates with a nitrogen mustard-like interstrand cross-link (*right panel*, cross-link shown in *red*) were produced as described under "Experimental Procedures." Schematics of products generated by the above enzymes are shown next to the corresponding substrates. *Red asterisks* indicate the positions of the ^{32}P-labeled nucleotides. Fragments invisible on the autoradiograph are shown in *gray*. *B–D*, incubation of the indicated substrates with FAN1 (*B*), EXO1 (*C*), or FEN1 (*D*). Aliquots were withdrawn after 0, 10 s, 2 min, 20 min, 40 min, and 80 min (*B*, *lanes 1–6* and *7–12*, respectively) or 0, 10 s, 1 min, 20 min, and 40 min (*C* and *D*, *lanes 1–6* and *7–12*, respectively). *Lowercase letters* on the *right* correspond to the products in *A*. The protein-to-DNA ratios were FAN1 (10:1), EXO1 (3.5:1), and FEN1 (1:2). *M*, low molecular weight marker (Affymetrix). The oligonucleotide sizes are indicated on the *left*. The *black asterisk* indicates the position of the non-cross-linked 37-mer oligonucleotide that was present in small amounts in the ICL substrate. The graphs below the autoradiographs of 20% denaturing polyacrylamide gels show the quantification of bands **a** or **b**, respectively. **b′** was generated by incisions in the top strand. *Error bars* show mean ± S.D. ($n = 3$). *E*, comparative analysis of cleavage products generated on the 3′ side of the cross-link by the three nucleases as well as by the nuclease-dead mutants of FAN1 (D960A) and EXO1 (D173N). The reaction conditions were as in Fig. 1*E*.

(Fig. 2*D*, *lanes 7–12*). This indicated that both enzymes made their primary endonucleolytic incisions 5′ from the ICL at position 18 of the labeled strand and that their exonuclease activity failed to traverse the cross-link. In contrast, incubation of the unmodified and the ICL substrate with FAN1 yielded similar products (Fig. 2*B*, *cf.* products **b** and **b′**, *lanes 1–6* and *lanes 7–12*), which arose from incisions in the labeled strand on the 3′ side of nucleotide 18 (and, therefore, 3′ from the cross-link in the ICL substrate) and their subsequent 5′ to 3′ degradation. Fig. 2*E* shows again that the activities were due to the enzymatic activities of FAN1 and EXO1 because their mutants failed to generate the specific products.

FAN1 Can Degrade a Linear Duplex Containing an ICL—Our experiments carried out with the 5′-labeled substrates showed that the flap endonuclease of FAN1 was relatively promiscuous because it incised the structures both 5′ and 3′ of the ss/ds boundary. Moreover, its activity did not appear to be influenced by the presence of the cross-link. Given that the enzyme is also a robust 5′ to 3′ exonuclease, as seen with the 3′-labeled substrates, we wanted to test whether the presence of a cross-link affected the latter activity. We therefore generated linear duplex substrates, either unmodified or containing a single ICL (Fig. 3*A*). As shown in Fig. 3, *B—D*, *lanes 1–6*, all three enzymes were able to degrade the unmodified 3′-labeled substrate, albeit inefficiently. Unexpectedly, FAN1 digestion gave rise to a similar range of short oligonucleotide products as those seen with the unmodified substrate (Fig. 3*B*, *cf. lanes 1–6* and *7–12*), which indicated that the enzyme traversed the ICL. This result confirms the nature of product **d′** in Fig. 1*B*. In contrast, EXO1 and FEN1 generated a series of products **b′** from the ICL substrate (Fig. 3, *C* and *D*, *lanes 7–12*), which indicated that the enzyme was arrested by the ICL. As shown in Fig. 3*E*, the inactive variants generated no products comparable with those generated by the wild-type proteins.

We then tested FAN1 (WT and D960A) activity on a linear ICL substrate with a recessed 5′ terminus. Also on this sub-

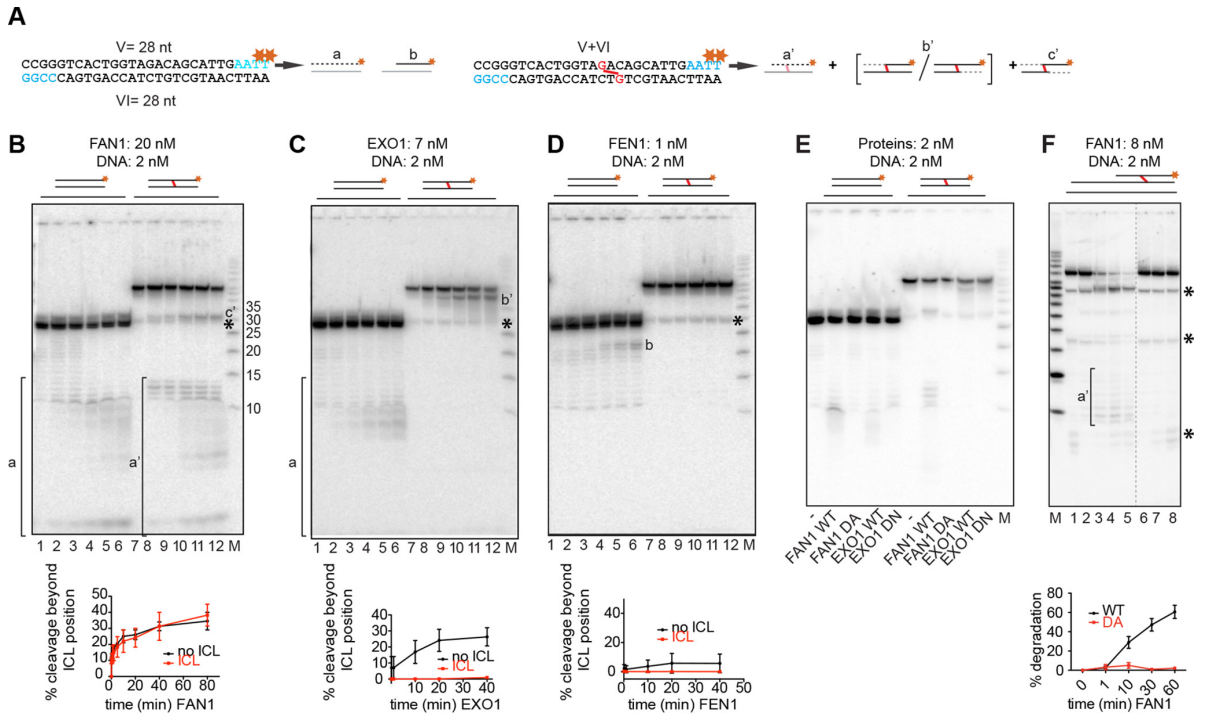

FIGURE 3. **FAN1 can also traverse a cross-link on a linear DNA substrate.** *A*, schematic of the linear DNA substrates labeled at the 3′ terminus of the upper strand by fill-in reactions with [α-^{32}P]dTTP. A schematic of the expected products is shown next to the corresponding substrates (**a**, **b**, and **a′**–**c′**). *Red asterisks* indicate the positions of the ^{32}P-labeled nucleotides. Fragments invisible on the autoradiograph are shown in *gray*. *B–D*, 5′ to 3′ exonuclease activities of FAN1 (*B*), EXO1 (*C*), and FEN1 (*D*) on the indicated substrates. Reaction conditions and protein concentrations were as in Fig. 2. *Lowercase letters* correspond to the products in *A*. *M*, low molecular weight marker (Affymetrix). The oligonucleotide sizes are indicated. The *black asterisk* indicates the position of the non-cross-linked 28-mer oligonucleotide that was present in small amounts in the ICL substrate. *Bottom panels*, quantifications of the degradation fragments. *Error bars* show mean ± S.D. (n = 3). *E*, comparative analysis of cleavage products generated on the 3′ side of the cross-link by the three nucleases as well as by the nuclease-dead mutants of FAN1 (D960A) and EXO1 (D173N). The reaction conditions were as in Fig. 1*E*. *F*, FAN1 WT and D960A activity on a recessed linear DNA substrate shown schematically above. Concentrations were as indicated, and samples were withdrawn after 1, 10, 30, or 60 min (WT, *lanes 2–5*) or after 1, 30, and 60 min (D960A, *lanes 6–8*). For the substrate sequence, see "Experimental Procedures." Quantifications are as indicated above. *Black asterisks* indicate the non-cross-linked oligonucleotides present in the substrate preparation. The *dashed line* represents missing lanes.

strate, the WT FAN1 protein was able to traverse the cross-link (Fig. 3*F*, *lanes 1–5* (WT) and *6–8* (D960A)).

FAN1 Requires Regions of ~10 Base Pairs on Either Side of the Flap for Cleavage—The experiments described above provided us with novel information concerning the biochemical properties of FAN1. Could these data help elucidate the role of FAN1 in ICL processing *in vivo*? That FAN1 is involved in ICL damage processing is beyond doubt, given that cells lacking the full complement of this protein are hypersensitive to cisplatin, MMC (7–10), and the nitrogen mustard BCNU (Fig. 5*E*). It is, however, unclear, which structures it addresses in genomic DNA and which products it generates. Evidence obtained from experiments carried out with an ICL-containing plasmid substrate and *X. laevis* egg extracts indicated that a replication fork encountering a cross-link stalls initially 20–40 base pairs from it because of the CDC45/MCM2–7/GINS replicative helicase remaining on the leading strand. This intermediate is not incised in this experimental system but persists until the arrival of the second fork from the opposite direction. Both helicase complexes are then unloaded, and the forks converge on the ICL. Unhooking was detected when the forks were just a single base pair from the cross-link (Ref. 25 and references therein). In this system, the nucleases primarily responsible for unhooking are believed to be SLX1/SLX4 on the 5′ side of the ICL and XPF/ERCC1 on the 3′ side (23). FAN1 has been shown not to be necessary (5, 23), and this finding is supported by our data, which suggest that FAN1 would not incise the X structure generated by the convergence of the two forks because it appears to require the ICL to be flanked by dsDNA on both sides of the flap, as witnessed by the fact that a Y junction is a poor FAN1 substrate (9) and that its exonuclease activity failed to generate fragments shorter than ~10 nucleotides (Fig. 2*B*). This suggests that FAN1 requires at least this length of dsDNA for binding.

We decided to test this hypothesis directly by generating two X-like substrates reminiscent of two converging replication forks separated by 12 (X-12) or eight (X-8) nucleotides (Fig. 4*A*). As shown in Fig. 4*B*, the former substrate was processed more efficiently than the latter by FAN1 WT but not by the D960A mutant. This difference was not caused by the lower stability of the X-8 substrate. Both the X-12 and X-8 substrates were annealed efficiently under the assay conditions, as shown by native PAGE (Fig. 4*B*, *bottom right panel*). Therefore, coupled with the finding that the enzyme processes Y structures with only limited efficiency (9), the above result confirms our hypothesis that FAN1 requires regions of ~10 base pairs on either side of the 5′ flap for cleavage.

FAN1-catalyzed Unhooking of ICLs

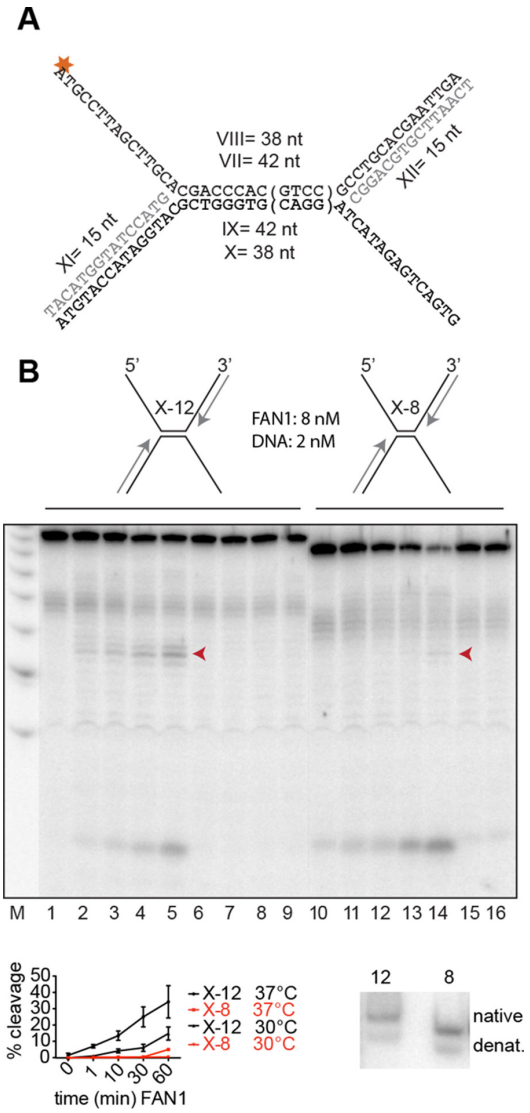

FIGURE 4. **FAN1 cleavage requires ~10 base pairs of dsDNA on both sides of the flap.** *A*, schematic of the X-12 and X-8 substrates. The tetranucleotide sequence in *parentheses* is absent from the X-8 substrate. The *asterisk* indicates the ^{32}P-labeled phosphate. *B*, the substrates were incubated with the enzymes at the indicated enzyme:substrate ratio at 37 °C for 1, 10, 30, and 60 min (FAN1-WT, *lanes 1–5* and *10–14*; FAN1 D960A, *lanes 6–9*) or 1 and 60 min (*lanes 15* and *16*). The X-8 substrate was extremely inefficiently processed even at 30 °C (*bottom left panel*), which indicates that the lack of flap cleavage was not caused by denaturation (*denat*) of the X structure. (Only the bands indicated by the *arrowheads* were quantified.) *Bottom right panel*, 10% native polyacrylamide gel of the two substrates, incubated at 37 °C for 30 min. This additional control experiment shows that both substrates were predominantly annealed under the conditions of the reaction.

MMC-induced DSBs Are Partially FAN1-dependent—If FAN1 is not involved in the resolution of replication forks converged at an ICL, then which structures does it address *in vivo*? This question is extremely difficult to answer, but it might be possible to gain some insights from the analysis of the products it generates from DNA treated with ICL-inducing agents. Processing of ICL-arrested replication forks by the FANC pathway (5) leads to lesion unhooking and to the generation of one or two DSBs that are subsequently repaired by homologous recombination (2). Some years ago, the *MUS81* gene was also implicated in ICL repair. Like *FAN1*, *MUS81* is not a known *FANC* gene. It encodes a protein that associates with EME1 to form a heterodimer that possesses 3′ flap endonuclease activity. Disruption of the *Mus81* locus in the mouse resulted in sensitivity to ICL-inducing agents and to a decreased number of MMC-induced DSBs in ES cells in S phase, as shown by pulsed field gel electrophoresis (26). This suggested that Mus81 might be involved in the conversion of ICLs to DSBs, which are in turn repaired by recombination (22). We decided to make use of this assay to learn whether we could find evidence of similar processing of ICL-containing DNA by FAN1. We treated U2OS (osteosarcoma) cells with siRNAs targeting either luciferase (negative control), FAN1, and/or MUS81 and, 2 days later, exposed them to MMC. Western blot analysis of extracts isolated 24 h after MMC treatment revealed that FAN1 and/or MUS81 knockdown was associated with a reduction in phosphorylation of replication protein A (RPA), CtIP, and H2AX (Fig. 5*A*), indicative of less extensive resection of DSBs. This was confirmed by an analysis of genomic DNA 18, 24, and 30 h after MMC treatment. As shown in Fig. 5*B*, the number of DSBs was reduced in FAN1 or MUS81 knockdown cells compared with the control. Surprisingly, quantification of the 24-h time point (Fig. 5*C*) revealed that knockdown of both MUS81 and FAN1 gave rise to a similar number of DSBs as the FAN1 knockdown alone. This effect was not caused by a proliferative defect of the siRNA-treated cells, as confirmed by EdU incorporation and FACS analysis (Fig. 5*D*). The finding that knockdown of FAN1 or MUS81 resulted in a similar sensitivity to MMC or carmustine and that knockdown of both mRNAs failed to cause additional sensitization (Fig. 5*E*) suggested that the two proteins might act together in the processing of a subset of ICL substrates.

Discussion

The current hypotheses of ICL processing (5) posits that the repair machinery has to deal with either one of two distinct replication-dependent lesions: a single fork arrested at the ICL (Fig. 6*A*) or an X-shaped structure that arises at the cross-link through the convergence of two forks or through replication fork traverse through the ICL (Fig. 6*B*). Our findings suggest that FAN1 would be able to process the single arrested fork alone by first releasing the flap and then unhooking/traversing the ICL to leave behind a short oligonucleotide attached to the leading strand template that could be bypassed by translesion polymerases (Fig. 6*C*). However, this scenario would apply solely if FAN1 were able to incise the flap when the ICL was more than four nucleotides 3′ from the ss/ds boundary. If it were closer, FAN1 would fail to release the flap. FAN1 could also initiate the processing of an X-shaped structure, by releasing the lagging strand flap and, therefore, forming a substrate for a second nuclease, such as MUS81, which could release the 3′ flap formed by the leading strand of the second fork. This mechanism would only work if the two forks were at least 15 base pairs apart, with the ICL being positioned more than four nucleotides from the left ss/ds boundary (Fig. 6*D*) to provide

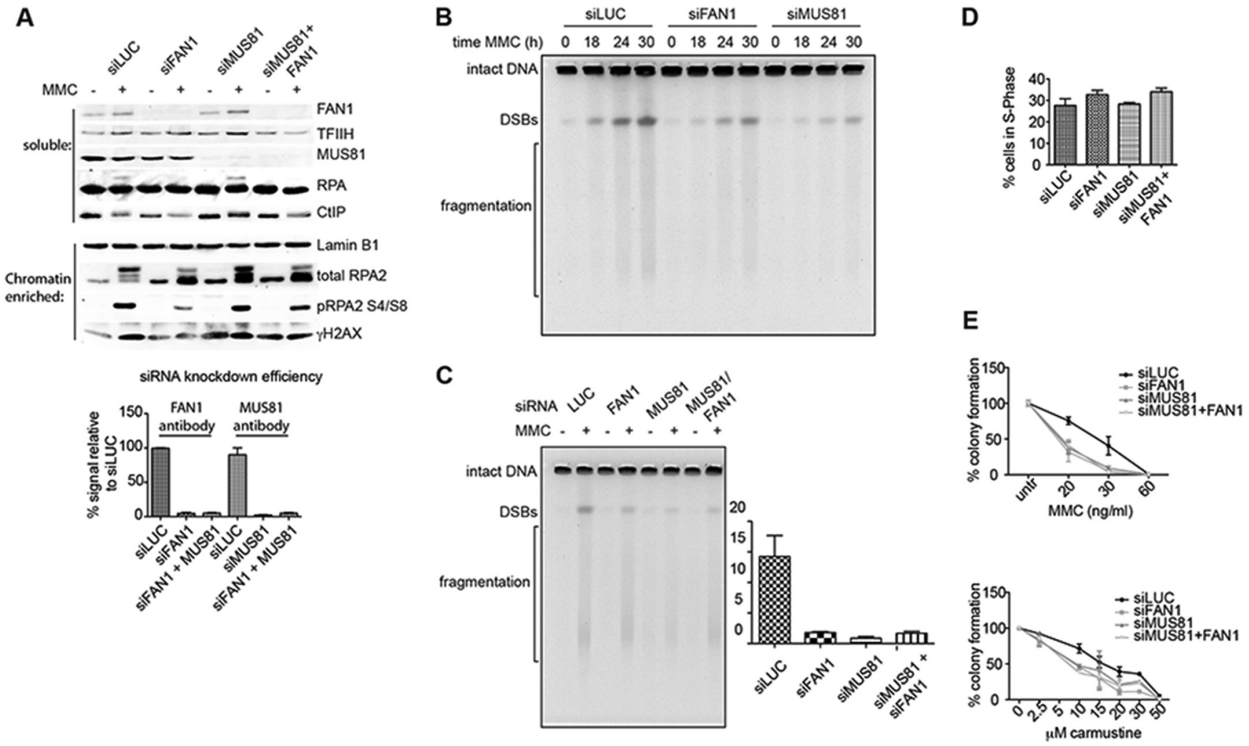

FIGURE 5. **FAN1- and/or MUS81-dependent DSB induction upon MMC treatment of human cells.** *A*, *top panel*, FAN1 and/or MUS81 siRNA-mediated knockdown efficiencies assessed by Western blotting of total cell extracts of untreated (−) or MMC-treated (+) U2OS cells. TFIIH was used as a loading control. Quantification of the knockdown efficiencies shown in the *bottom panel* was carried out using ImageJ, and the graph was produced by GraphPad Prism (n = 3). The same extracts were probed for the markers of DSB metabolism RPA and CtIP. *Center panel*, Western blot analysis of the chromatin-enriched fraction of the above extracts probed for RPA, phospho-RPA, and γH2AX. Lamin was used as the loading control. *B*, time course of DSB formation assessed by pulsed field gel electrophoresis after MMC treatment (3 μg/ml) of U2OS cells, in which FAN1 or MUS81 were knocked down by siRNA. *C*, representative pulsed field gel electrophoresis image of DSBs induced by 24 h MMC (3 μg/ml) treatment of U2OS cells in which FAN1 and/or MUS81 were knocked down. The *left panel* shows a quantification of three independent experiments. siLUC was used as a control, and the ratio of DSBs of the MMC-treated samples divided by the untreated samples is shown for each siRNA condition. *Error bars* show mean ± S.E. *D*, quantifications of a FACS analysis of EdU-labeled cells pretreated with the indicated siRNAs and subsequently treated for 24 h with 3 μg/ml MMC. The knockdown did not affect cell viability during the course of the experiment. *Error bars* show mean ± S.E. (n = 3). *E*, clonogenic survival assay of U2OS cells treated with the indicated siRNAs and drugs. Colonies were counted 8 days after treatment, and MMC was washed out 24 h after treatment.

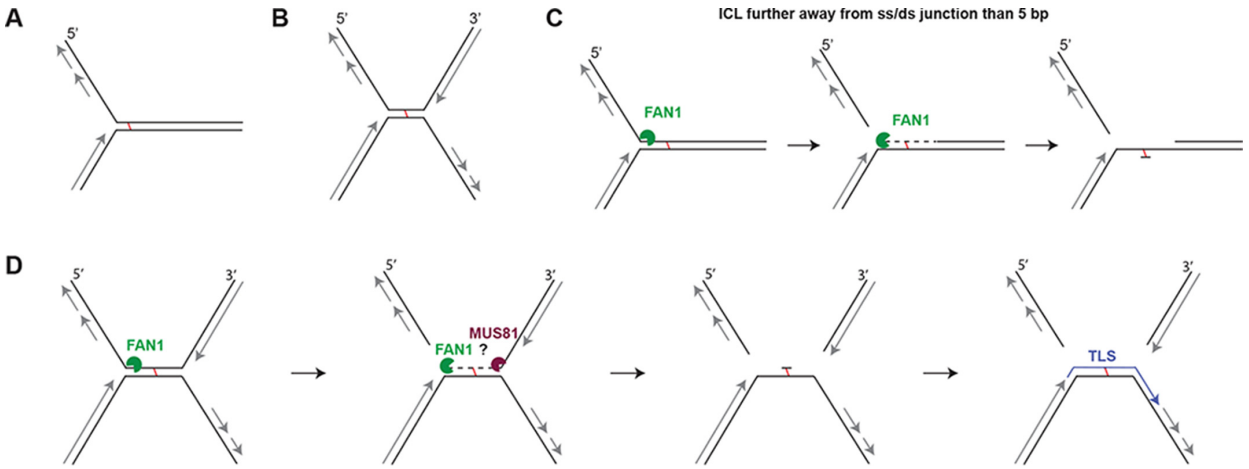

FIGURE 6. **The putative mechanism of FAN1-dependent ICL unhooking.** *A*, schematic of a single replication fork arrested at an ICL. *B*, the X-shaped structure arising through the convergence of two replication forks at an ICL. *C*, a single fork arrested at an ICL that is more than 5 bp from the ss/ds boundary would be incised by FAN1 5′ from the ICL, which would release the lagging strand. The enzyme would then degrade the nicked strand in a 5′-to-3′ direction to generate a substrate for translesion polymerases and subsequent repair by nucleotide excision repair. *D*, an X-shaped structure that contains a duplex longer than 10 bp where the lagging and leading strand flaps could be released by the action of FAN1 and, *e.g.*, MUS81.

FAN1-catalyzed Unhooking of ICLs

FAN1 with a duplex platform to which it has to bind on either side of the flap (see above).

The hypothesis of ICL processing favoring the convergence of two forks arriving from opposite directions has recently received additional support (25). Of the known 5′ flap endonucleases, SLX1/SLX4 is emerging as the strongest candidate for introducing the first incision, with XPF/ERCC1 being the prime candidate for the 3′ incision. The involvement of FAN1 in the processing of a subset of these structures, where the forks remained some distance apart because of topological restraints, for example, cannot be excluded. But FAN1 does not necessarily have to be involved in ICL processing. We originally proposed that the enzyme might be involved downstream of the unhooking step, possibly during homologous recombination, because the enzyme processed a D loop structure with very high efficiency and because foci of γ-H2AX, RPA, and RAD51, which are generally believed to be markers of DSBs, appeared with similar kinetics in FAN1-proficient and knockdown cells but persisted longer in depleted cells (9), and depletion of MUS81 results in a similar phenotype (27). Interestingly, these foci also form after replication fork stalling (22, 28–30), and FAN1 has been reported to accumulate at replication forks stalled by aphidicolin (31). It is therefore possible that the persistent foci of γ-H2AX, RPA, and RAD51 represent stalled replication forks that collapsed because of the deficiency in enzymes able to mediate their restart and that are being processed by an alternative, less efficient mechanism.

Recently, the crystal structure of *Pseudomonas aeruginosa* FAN1 bound to a 5′ flap (32) revealed that the enzyme (lacking the ubiquitin-binding zinc finger domain) bends the DNA at the flap position by extensively interacting with the dsDNA regions on either side of the flap and that it cleaves the fourth phosphodiester moiety 3′ from the ss/dsDNA boundary. These observations fully agree with our findings. When our manuscript was in the final stages of preparation, two additional publications described the crystal structure of human FAN1 (also lacking the ubiquitin-binding zinc finger domain) bound to different flap substrates. In one study (33), FAN1 was shown to bind with a much greater affinity to substrates containing a 5′-phosphorylated flap of only one or two nucleotides, which it then incised in exonucleolytic steps of three nucleotides. In this way, it was able to unhook a triazole cross-link in a way analogous to that described in our study. The second study deployed a series of substrates in which the position of the ICL was moved further away from the ss/ds boundary of the flap (34). In the latter work, cleavage efficiency was seen to increase as the cross-link was moved 6, 12, or 16 base pairs away from the ss/ds boundary. This work predicted that, on the latter substrates, the enzyme would successfully traverse the ICL. In this scenario, it would be able to unhook the cross-link without the assistance of other enzymes. Taken together, our study and those of Wang *et al.* (33) and Zhao *et al.* (34) demonstrate that FAN1 can cleave long 5′ flaps such as those that would arise at blocked replication forks but that the enzyme may also process and unhook short flaps and nicks generated at ICLs by other nucleases that may not be able to bypass the ICL. In the latter reaction, FAN1 may be partially redundant with SNM1A, as is the case in *Schizosaccharomyces pombe* (35). Like FAN1, the latter enzyme has the ability to traverse cross-links, providing that a single-stranded nick 5′ from the ICL is present in the vicinity (36).

Because ICLs block the progression of transcription and replication machineries and because they can bring about different distortions in the helical structure of DNA (37), it is not surprising that nature has evolved several mechanisms of dealing with these extremely cytotoxic lesions. The investigation of which enzyme or pathway acts when and on which substrate(s) must await the outcome of genetic experiments that are currently in progress in several laboratories.

Author Contributions—J. P. helped design and carried out the experiments and cowrote the manuscript. S. M. synthesized the ICL oligonucleotides; O. D. S. helped design the experiments; and J. J. conceived the study, helped design the experiments, and wrote the manuscript.

Acknowledgments—We thank Anja Saxer for technical help. We also thank Stephanie Bregenhorn for the purified recombinant EXO1, Pavel Janscak for the MUS81 antibody, Alessandro Sartori for the CtIP antibody, and Svenja Kaden for helpful discussions and critical reading of the manuscript.

References

1. Schärer, O. D. (2005) DNA interstrand crosslinks: natural and drug-induced DNA adducts that induce unique cellular responses. *ChemBioChem* **6,** 27–32
2. Hinz, J. M. (2010) Role of homologous recombination in DNA interstrand crosslink repair. *Environ. Mol. Mutagen.* **51,** 582–603
3. Auerbach, A. D. (2009) Fanconi anemia and its diagnosis. *Mutat. Res.* **668,** 4–10
4. Crossan, G. P., and Patel, K. J. (2012) The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. *J. Pathol.* **226,** 326–337
5. Zhang, J., and Walter, J. C. (2014) Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair* **19,** 135–142
6. Clauson, C., Schärer, O. D., and Niedernhofer, L. (2013) Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb. Perspect. Biol.* **5,** a012732
7. Liu, T., Ghosal, G., Yuan, J., Chen, J., and Huang, J. (2010) FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science* **329,** 693–696
8. Smogorzewska, A., Desetty, R., Saito, T. T., Schlabach, M., Lach, F. P., Sowa, M. E., Clark, A. B., Kunkel, T. A., Harper, J. W., Colaiácovo, M. P., and Elledge, S. J. (2010) A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol. Cell* **39,** 36–47
9. Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavó, E., Sartori, A. A., Hengartner, M. O., and Jiricny, J. (2010) Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell* **142,** 77–88
10. MacKay, C., Déclais, A. C., Lundin, C., Agostinho, A., Deans, A. J., MacArtney, T. J., Hofmann, K., Gartner, A., West, S. C., Helleday, T., Lilley, D. M., and Rouse, J. (2010) Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell* **142,** 65–76
11. Pennell, S., Déclais, A. C., Li, J., Haire, L. F., Berg, W., Saldanha, J. W., Taylor, I. A., Rouse, J., Lilley, D. M., and Smerdon, S. J. (2014) FAN1 activity on asymmetric repair intermediates is mediated by an atypical monomeric virus-type replication-repair nuclease domain. *Cell Rep.* **8,** 84–93
12. Munoz, I. M., Szyniarowski, P., Toth, R., Rouse, J., and Lachaud, C. (2014) Improved genome editing in human cell lines using the CRISPR method. *PloS ONE* **9,** e109752
13. Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka,

H., Horimoto, S., Takeda, S., and Jiricny, J. (2010) KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 21553–21557
14. Zhou, W., Otto, E. A., Cluckey, A., Airik, R., Hurd, T. W., Chaki, M., Diaz, K., Lach, F. P., Bennett, G. R., Gee, H. Y., Ghosh, A. K., Natarajan, S., Thongthip, S., Veturi, U., Allen, S. J., Janssen, S., Ramaswami, G., Dixon, J., Burkhalter, F., Spoendlin, M., Moch, H., Mihatsch, M. J., Verine, J., Reade, R., Soliman, H., Godin, M., Kiss, D., Monga, G., Mazzucco, G., Amann, K., Artunc, F., Newland, R. C., Wiech, T., Zschiedrich, S., Huber, T. B., Friedl, A., Slaats, G. G., Joles, J. A., Goldschmeding, R., Washburn, J., Giles, R. H., Levy, S., Smogorzewska, A., and Hildebrandt, F. (2012) FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nat. Genet.* **44,** 910–915
15. Orans, J., McSweeney, E. A., Iyer, R. R., Hast, M. A., Hellinga, H. W., Modrich, P., and Beese, L. S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145,** 212–223
16. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A., and Tainer, J. A. (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145,** 198–211
17. Tsutakawa, S. E., Lafrance-Vanasse, J., and Tainer, J. A. (2014) The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. *DNA Repair* **19,** 95–107
18. Mukherjee, S., Guainazzi, A., and Schärer, O. D. (2014) Synthesis of structurally diverse major groove DNA interstrand crosslinks using three different aldehyde precursors. *Nucleic Acids Res.* **42,** 7429–7435
19. Guainazzi, A., Campbell, A. J., Angelov, T., Simmerling, C., and Schärer, O. D. (2010) Synthesis and molecular modeling of a nitrogen mustard DNA interstrand crosslink. *Chemistry* **16,** 12100–12103
20. Angelov, T., Guainazzi, A., and Schärer, O. D. (2009) Generation of DNA interstrand cross-links by post-synthetic reductive amination. *Org. Lett.* **11,** 661–664
21. Bregenhorn, S., and Jiricny, J. (2014) Biochemical characterization of a cancer-associated E109K missense variant of human exonuclease 1. *Nucleic Acids Res.* **42,** 7096–7103
22. Hanada, K., Budzowska, M., Davies, S. L., van Drunen, E., Onizawa, H., Beverloo, H. B., Maas, A., Essers, J., Hickson, I. D., and Kanaar, R. (2007) The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nat. Struct. Mol. Biol.* **14,** 1096–1104
23. Klein Douwel, D., Boonen, R. A., Long, D. T., Szypowska, A. A., Räschle, M., Walter, J. C., and Knipscheer, P. (2014) XPF-ERCC1 acts in unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol. Cell* **54,** 460–471
24. Harrington, J. J., and Lieber, M. R. (1994) Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. *Genes Dev.* **8,** 1344–1355
25. Zhang, J., Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., and Walter, J. C. (2015) DNA interstrand cross-link repair requires replication-fork convergence. *Nat. Struct. Mol. Biol.* **22,** 242–247
26. Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar, R. (2006) The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. *EMBO J.* **25,** 4921–4932
27. Nair, N., Castor, D., Macartney, T., and Rouse, J. (2014) Identification and characterization of MUS81 point mutations that abolish interaction with the SLX4 scaffold protein. *DNA Repair* **24,** 131–137
28. Ward, I. M., and Chen, J. (2001) Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *J. Biol. Chem.* **276,** 47759–47762
29. Vassin, V. M., Anantha, R. W., Sokolova, E., Kanner, S., and Borowiec, J. A. (2009) Human RPA phosphorylation by ATR stimulates DNA synthesis and prevents ssDNA accumulation during DNA-replication stress. *J. Cell Sci.* **122,** 4070–4080
30. Petermann, E., Orta, M. L., Issaeva, N., Schultz, N., and Helleday, T. (2010) Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol. Cell* **37,** 492–502
31. Chaudhury, I., Stroik, D. R., and Sobeck, A. (2014) FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Mol. Cell. Biol.* **34,** 3939–3954
32. Gwon, G. H., Kim, Y., Liu, Y., Watson, A. T., Jo, A., Etheridge, T. J., Yuan, F., Zhang, Y., Kim, Y., Carr, A. M., and Cho, Y. (2014) Crystal structure of a Fanconi anemia-associated nuclease homolog bound to 5′ flap DNA: basis of interstrand cross-link repair by FAN1. *Genes Dev.* **28,** 2276–2290
33. Wang, R., Persky, N. S., Yoo, B., Ouerfelli, O., Smogorzewska, A., Elledge, S. J., and Pavletich, N. P. (2014) DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science* **346,** 1127–1130
34. Zhao, Q., Xue, X., Longerich, S., Sung, P., and Xiong, Y. (2014) Structural insights into 5′ flap DNA unwinding and incision by the human FAN1 dimer. *Nat. Commun.* **5,** 5726
35. Fontebasso, Y., Etheridge, T. J., Oliver, A. W., Murray, J. M., and Carr, A. M. (2013) The conserved Fanconi anemia nuclease Fan1 and the SUMO E3 ligase Pli1 act in two novel Pso2-independent pathways of DNA interstrand crosslink repair in yeast. *DNA Repair* **12,** 1011–1023
36. Wang, A. T., Sengerova, B., Cattell, E., Inagawa, T., Hartley, J. M., Kiakos, K., Burgess-Brown, N. A., Swift, L. P., Enzlin, J. H., Schofield, C. J., Gileadi, O., Hartley, J. A., and McHugh, P. J. (2011) Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes Dev.* **25,** 1859–1870
37. Muniandy, P. A., Liu, J., Majumdar, A., Liu, S. T., and Seidman, M. M. (2010) DNA interstrand crosslink repair in mammalian cells: step by step. *Crit. Rev. Biochem. Mol. Biol.* **45,** 23–49

2.2 Additional Observations

2.2.1 5' Flap cleavage activity in human TSCRE2 extracts

FAN1 was the only one of the three tested endonucleases able to perform the unhooking reaction *in vitro* under the used conditions by cleaving 5' and 3' of the lesion. However, FEN1 and EXO1 were also able to efficiently cleave the substrate 5' to the ICL, so we wondered what the relative contribution of these enzymes to ICL processing in cells might be. By using nuclear or whole cell extracts, we aimed to establish which nuclease has a higher propensity to cleave flap structures, either in general or in the vicinity of an ICL. The extracts should reflect the situation in the cell, with respect to the relative abundance of the different proteins, the presence of potential interaction partners and cofactors, and to a certain extent their relative activities. We used TSCRE2 cells, which are derivatives of the TK6 human lymphoblastoid cell line. They were either WT or knockout (KO) for *FAN1*, *EXO1* or both. Using the 3' labeled substrates, we could show that the efficiency of endonucleolytic cleavage was not impaired in the nuclear extracts derived from any of the cell lines used (see **Figure R1**). This means that the major flap nuclease activity in the extracts is neither FAN1, nor EXO1. A good candidate for the protein performing the cleavage in the extracts could be FEN1, as no cleavages were made 3' to the ICL and the cross-link was only partially unhooked on the 5' side. This reflects FEN1 rather than FAN1 substrate specificity. A similar cleavage pattern was seen in nuclear extracts of WT HEK293 cells. As *FEN1* deletion is lethal in human cells, we switched to the chicken DT40 cell line, where *FEN1* KO is viable.

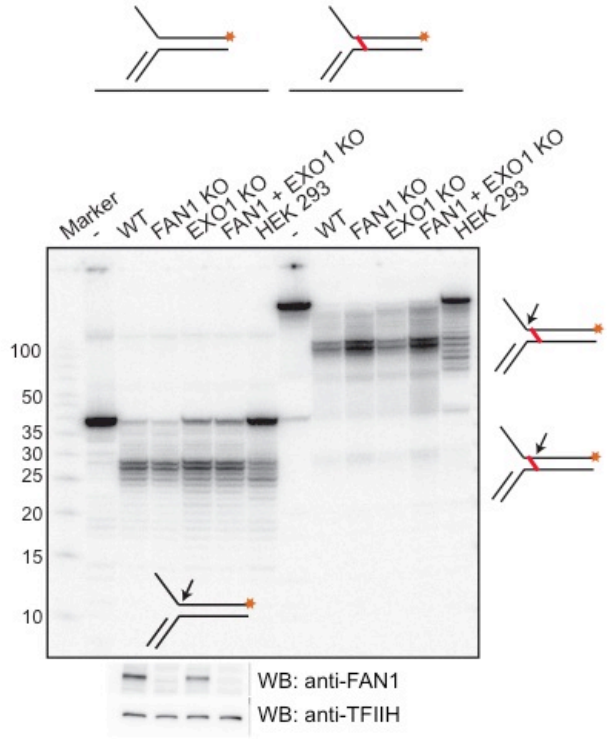

Figure R1: Nuclease assays with TSCER2 nuclear extracts on 5' flap substrates with or without an ICL. Nuclear extracts from WT, FAN1, EXO1 and double KO TSCER2 cells were incubated with the depicted flap substrates, radioactively labeled at the 3' end (orange star). The products were separated on a 20% denaturing polyacrylamide gel. The ICL is indicated by a red line in the drawing. HEK293 nuclear extracts were used as an additional control. The cleavage positions are indicated by arrows on the drawings (below for the non-cross-linked and on the right for the cross-linked substrates). The western blot (WB) shown below the DNA gel of WT, FAN1, EXO1, or double KO extracts confirms the absence of FAN1. TFIIH was used as a loading control. Marker: Low molecular weight marker (Affymetrix), oligonucleotide sizes are indicated on the left.

2.2.2 FEN1 is responsible for almost all 5' flap cleavage in chicken cell extracts, but FAN1 activity becomes detectable upon overexpression

As FEN1 KO is lethal in human cells but viable in DT40 chicken cells, we made use of FEN1 KO cells in the DT40 system that were generated by Dr. Saho Kobayashi in our laboratory. To test whether FEN1 was really the major 5' flap endonuclease in the extracts, we performed *in vitro* nuclease assays on 5' flap DNA structures using extracts of WT, FAN1 KO or FEN1 KO DT40 cells. FAN1 KO cell extracts cleaved the substrate in a very similar way as the extracts derived from the WT cells, which was consistent with the results using the human TSCER2 cell line extracts. FEN1 KO, on the other hand, almost completely lacked the main cleavage activity (**Figure R2A**). This means that FEN1 is indeed the most efficient 5' flap endonuclease present in extracts derived from these cells and that FAN1 (or any other endonuclease) activity is minimal under the used conditions.

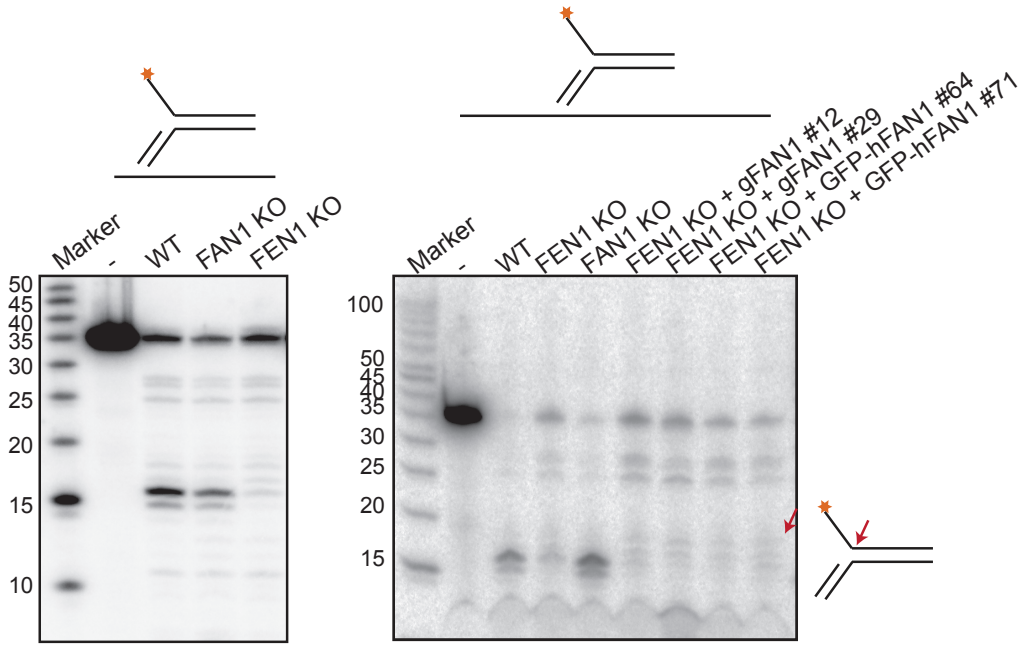

Figure R2: Activity of Chicken DT40 extracts on 5' flap structures. A) Extracts from WT, FAN1 KO, and FEN1 KO DT40 cells were used in nuclease assays. The 5' flap substrate, as depicted, was labeled at the 5' end (orange star). Marker and gel as in Figure R1. Oligonucleotide sizes are indicated on the left. **B)** The same type of assay as in A) was performed using extracts from FEN1 KO cells transfected with a plasmid expressing either chicken (g)FAN1 or human (h)FAN1. The numbers indicate different clones. The red arrow indicates the band corresponding to the cleavage position as shown on the schematic drawing on the right that appears specifically upon FAN1 overexpression.

As FAN1 activity was very low in the extracts used for the nuclease assays, we aimed to see whether we could detect FAN1-dependent cleavage when overexpressing FAN1 in the chicken cells. To this end, we performed the same type of experiment using DT40 cells that contained a FAN1-expressing plasmid. Two clones overexpressing human FAN1 (hFAN1, #64 and #71) and two overexpressing chicken FAN1 (gFAN1, #12 and #29) were selected. As shown in **Figure R2B**, a weak 18-mer band appeared upon incubation of the DNA substrate with these extracts, as opposed to the 16-mer product generated in WT or FAN1 KO extracts. Based on its size and on its presence solely in reactions using extracts of FAN1 overexpressing cells, we concluded that FAN1 was the responsible endonuclease giving rise to this product. This means that FAN1 could potentially cleave also in the extracts, but that its abundance or activity is not high enough in WT extracts and therefore it is masked by FEN1 activity. For efficient processing of DNA molecules by FAN1 in the cell, there have to be some regulatory mechanisms that keep FEN1 away and provide FAN1 with access to the correct DNA metabolic pathway. This regulation could be of any type, ranging from accessibility due to nuclear localization, post-translational modifications, cell-cycle dependent fluctuations, accessory proteins or cofactors etc.

2.2.3 FEN1 depletion sensitizes chicken cells to MMC and this effect can be rescued by FAN1 overexpression

This experiment was performed together with Dr. Saho Kobayashi. We have seen from the experiments listed above that FEN1 has the highest activity in the extracts under our experimental conditions. As FEN1 was also able to incise DNA 5' to an ICL at a model stalled replication fork (see **chapter 2.1**), we wondered whether it could also be involved in the unhooking step of ICL repair, or whether it is somehow actively kept away from these structures. As mentioned in the introduction, **chapter 1.4**, FEN1 needs to thread the free 5' end through a hole in its tertiary structure for flap recognition and efficient cleavage. This requirement would be somewhat contradictory to a function in ICL unhooking and the recovery of stalled replication forks, where no free 5' end is available. However, as also mentioned before, FEN1 has in addition to its 5' flap endonuclease also a gap endonuclease activity, which allows it to cleave at the base of a model stalled replication fork (152). This GEN activity could allow it to unhook ICLs by cleaving a replication fork stalled at an ICL. Indeed, FEN1 KO DT40 cells were hypersensitive to MMC compared to the WT cells (**Figure R3A**). The question then arose whether the sensitivity is specific to an unhooking function, as is expected for FAN1 or whether it is involved in a step other than the FAN1-dependent one. To address this question, we used the FEN1 KO DT40 cells overexpressing FAN1 to see whether this rescues the sensitivity, which would mean that the two endonucleases have the ability to cleave similar ICL-dependent substrates in the cell. Indeed, the overexpression of either chicken or human FAN1 in the FEN1 KO cells led to a rescue of sensitivity towards MMC (**Figure R3A**). Clone number 29 for gFAN1 and clone number 64 for hFAN1 expressed higher GFP levels compared to clones 12 (gFAN1) and 71 (hFAN1) as seen in the FACS profiles (**Figure R3B**). Interestingly, the overexpression also correlated with the extent of resistance in the survival study, which indicates that FAN1 is most likely the responsible factor for the change in sensitivity. This could be explained by the fact that the activity and/ or abundance of FAN1 is quite low in the cells, as pointed out in the previous section, and that it is only able to perform a subset of incisions in the absence of FEN1, which has a partially redundant role. Upon overexpression of FAN1, however, there is enough enzyme to fully rescue the phenotype. We are currently generating the double KO cell line and, given that this is not lethal in DT40, would like to carry out a similar analysis on these cells, where we would expect a further decrease in survival compared to the single KO cell lines.

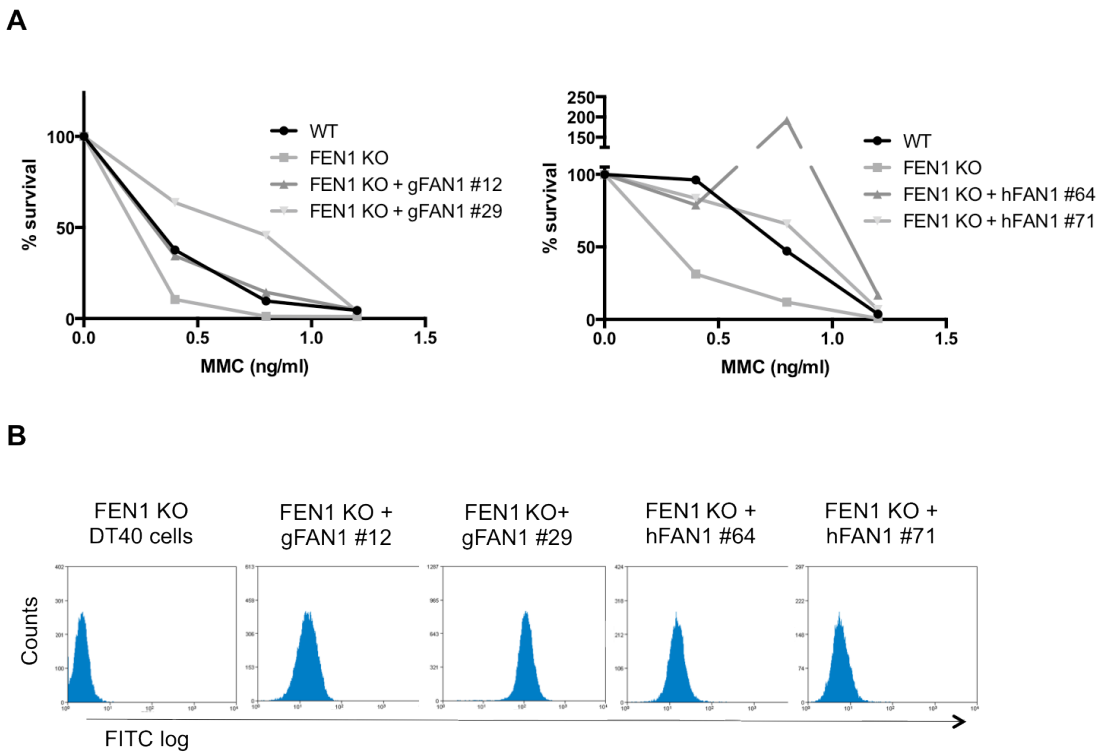

Figure R3: Sensitivity of chicken DT40 cells to MMC. A) Survival of FEN1 KO DT40 cells and two clones each overexpressing chicken (g, left graph) or human (h, right graph) FAN1 compared to WT cells after MMC treatment at the indicated concentrations. Survival was assessed by colony formation in methylcellulose-containing medium. Numbers indicate the different clones. **B)** FACS profiles of the cells used in A). Counts (y-axis) are plotted against FITC log (x-axis) for the GFP signal that is expressed from the same vector as gFAN1 or in the same open reading frame as hFAN1.

2.2.4 FAN1 protein levels in human cells are higher in the absence of FEN1 protein

As FEN1 seems to outcompete FAN1 under normal conditions, we wondered whether the two proteins could be alternatively regulated, i.e. whether FAN1 could become more abundant upon the loss of FEN1. Furthermore, we wanted to see whether the results obtained with the DT40 cell lines could also be relevant for human cells. First, we performed a clonogenic survival assay with U2OS cells treated with siRNA against FAN1, FEN1 or both proteins together (**Figure R4A**). As could be expected from the results in DT40 cells, the lack of both proteins led to a decrease in survival upon treatment with MMC and the co-depletion increased the effect further. This could indeed indicate that the two proteins act on a similar substrate in the process of ICL repair and that they could partially compensate for each other's loss. To see whether this compensation mechanism could even be visible on the protein or mRNA level, we prepared chromatin and whole cell extracts from U2OS cells and studied these levels upon depletion of one of the two proteins and after treatment with MMC. Whereas FEN1 expression or stability did not seem to be affected by the absence of FAN1 (but was rather lower at mRNA level), FAN1 levels were slightly higher upon FEN1 KO both,

at the protein and mRNA levels (**Figure R4B/C**). Similar results were obtained for protein levels in HEK293 cells (**Figure R4B**). This is consistent with the previous results, showing that FAN1 could compensate for FEN1 loss and it might do so by overexpression or an increase in stability. Based on its high activity in extracts, it also makes sense that FEN1 levels are not increased upon FAN1 downregulation, as it is probably already saturated under normal conditions. The increase of FAN1 mRNA levels upon FEN1 knockdown would have to be confirmed further, but if it is true, this could mean that the increase in intermediate DNA structures normally cleaved by FEN1 or FAN1 upon the loss of the former leads to a signaling cascade which ultimately triggers transcription of the *FAN1* gene. The fact that the levels are increased in all, the chromatin, soluble and whole cell fractions indicates that this is not due to an increased recruitment or retention of proteins on the chromatin, but rather to a global response in stability or expression level.

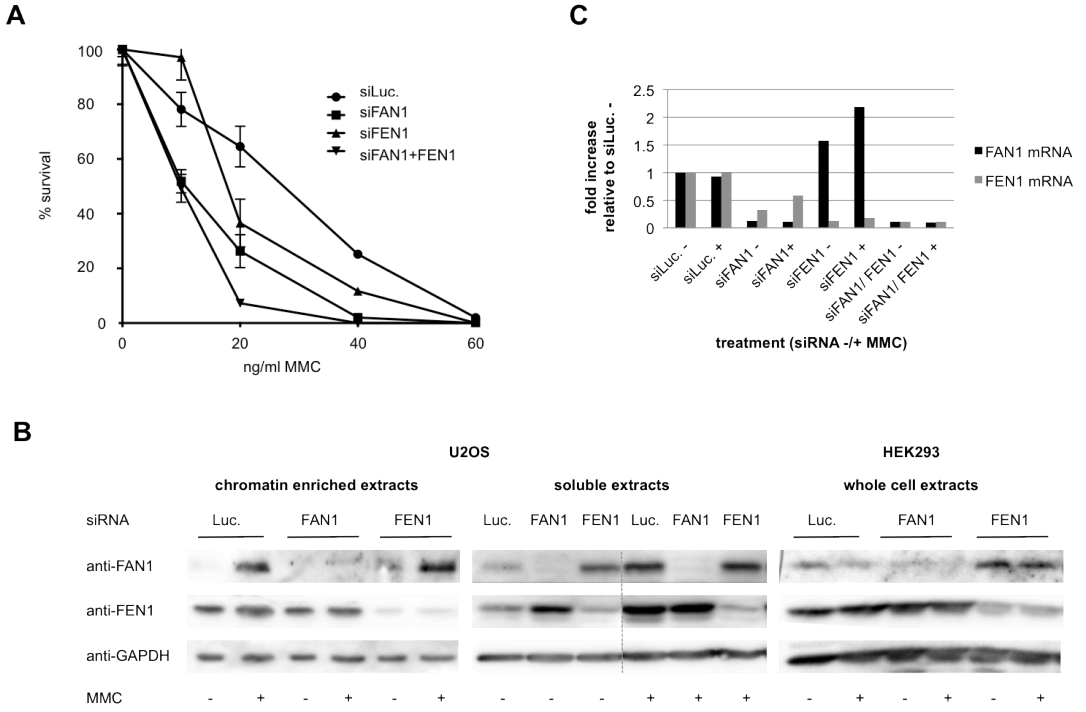

Figure R4: Human cells lacking FAN1, FEN1, or both. A) U2OS cells were treated with the indicated siRNAs and doses of MMC and survival was assessed by clonogenic survival assay. **B)** Western blot analysis of U2OS chromatin-enriched or soluble protein extracts or HEK293 whole cell extracts treated with the indicated siRNAs and MMC (1 µg/ml, 24 h). GAPDH was used as a loading control. **C)** mRNA levels as assessed by quantitative reverse-transcription PCR of U2OS cells treated as indicated in B).

3 Discussion

3.1 *In vitro* unhooking activity

FAN1 has been proposed to be important for the process of ICL repair and was already shown to have both endo- and exonucleolytic activities *in vitro*. By using DNA substrates containing site-specific ICLs, I could show that FAN1 is able to make incisions both 5' and 3' of the lesion under the conditions used and that it proceeds in an exonucleolytic manner once it passed the site of the ICL. This was true for both 5' flapped as well as linear double-stranded DNA, indicating that FAN1's ability to process DNA molecules containing a cross-link is not dependent on its 5' flap endonuclease activity. FEN1 and EXO1, two other endo- and exonucleases were not able to perform a similar reaction; they were stalled by the presence of the cross-link and could not perform the unhooking reaction on both sides of it. Neither FEN1 nor EXO1 have been implicated in ICL repair so far and this would be in line with their inability to unhook cross-links *in vitro*. When looking at the crystal structures of FAN1, FEN1 and EXO1 bound to DNA, it is intriguing to note that FEN1 and EXO1 use a two-nucleotide fraying mechanism to position the scissile phosphate at their active site (150,175). This means that the two terminal bps of the 5' flapped strand are unpaired to allow bending of the ssDNA towards the active site, divalent metal ion binding and cleavage. As such an unpairing mechanism is not possible if a cross-link is present at the base of the junction, it could explain why the FEN1 superfamily of endonucleases, to which both EXO1 and FEN1 belong, are able to make the incision directly at the junction, but not further downstream. FAN1 on the other hand does not seem to involve such an unpairing and bending mechanism for recognition and cleavage and makes only relatively loose contacts with the phosphates of the second and third nucleotide after the junction within the DNA strand it is cleaving. This could explain its ability to process highly-distorted DNA structures (166). Only Zhao et al. proposed a substrate unwinding mechanism for FAN1 nuclease activity from their crystal structure, but this is in contrast to the other published FAN1 structures (169).

Close examination of the biochemical activity of FAN1 reveals that its exonuclease activity is unusual, for example in comparison with EXO1. It is much more active at the entry point (as revealed by the strong bands on the gels) and loses its activity upon travelling away from the entry point. Furthermore, it seems to make stronger incisions at every third nucleotide. Together with the fact that it has only one active site, the mutation of which abolishes both activities, this indicates that the apparent exonuclease activity is rather a processive endonuclease cleaving always at a distance of 3 nucleotides from the entry site. FAN1 endonuclease usually makes one major cut at the preferred cleavage site and one (or more) minor ones on each side of it, probably due to some flexibility in the active site. Thus, when looking at the degradation of the 3' labeled substrates, one observes an exonuclease-like pattern. This mode of action might also be the reason why FAN1 is not blocked by the presence of the ICL as opposed to EXO1 for example. A similar observation was made by

Wang et al., who showed that (truncated) hFAN1 has a relatively flexible active site leading to a ±1 nt shift in the scissile phosphate register (168).

What these authors also provided was a direct proof of unhooking of a triazole ICL by analyzing the FAN1-generated products by LC-MS. They also demonstrated the need for a 5' phosphate that makes contacts with FAN1. Hence, shorter flaps were cleaved more efficiently than long ones. Furthermore, the addition of a 3' flap on the nicked strand enhanced activity. The nicked substrate without a 5' flap, or those with 1- or 2 nt flaps were cleaved efficiently under physiological salt conditions (150 mM), whereas longer flaps needed lower ionic strength. I observed efficient cleavage of all flap lengths used, although I have not used substrates described in their studies. Moreover, I always worked at low salt concentrations. It is unclear why changing the salt concentration results in such a shift in activity and substrate requirements, but given the fact that the cellular salt concentration is around 150 mM, it could be assumed that FAN1 in the cell has a more restricted substrate spectrum than what we see *in vitro* and that it actually prefers shorter flaps or even nicked substrates over long flaps.

It would be interesting to see whether FAN1 possesses residues that affect its ability to process ICLs to a greater extent than undamaged substrates. Such a separation-of-function mutant would be a useful tool to study FAN1's functions in the cell within and outside of ICL repair. According to hFAN1's crystal structure, the 5' phosphate makes contacts with a highly basic pocket at the C-terminal domain of FAN1, while the next two phosphodiester groups make relatively few contacts and only the nonbridging oxygen atom of the third phosphodiester group coordinates the metal ion at the active site (168). This explains the cleavage by FAN1 in 3-nt spacings and might explain its ability to accomodate and cleave distorted and cross-linked DNA substrates. One way of creating a FAN1 version that would exclusively cleave unmodified DNA molecules would be by narrowing the active center around the second and third phosphodiester groups and tightening the contacts with them. This would have to be done by randomly mutating the residues around that region, as it is not easy to predict the effects of these mutations beforehand.

3.2 *In vivo* implications of FAN1 unhooking activity

DSB formation was decreased in the absence of FAN1 in U2OS cell lines upon treatment with MMC, which could argue in favor of a function in unhooking. However, the occurrence of these DSBs is not a direct proof of unhooking and it could be that they disappear because of a more efficient repair of the lesion in the absence of FAN1 (decreasing the intermediate DSBs faster). This is however unlikely because of the increase in DSBs over time (see **Figure 5B**) and the notion that FAN1 depletion leads to slower DSB repair, as exemplified by the disappearance of foci of DSB repair markers such as RPA, RAD51 or γH2AX (158,159). It might seem contradictory that DSB induction is decreased upon FAN1 loss, but these markers persist longer. However, the type of damage was not strictly comparable, as the PFGE experiments were done with MMC, which means that the DSBs were induced

specifically as an intermediate step of ICL repair. The amount of MMC used was not compatible with survival, but it led to a high load of ICLs and repair intermediates, making the effects better visible. The DSBs seen by PFGE are not a direct sign of cell death or survival upon MMC treatment, as FAN1 depleted cells have a lower DSB load but are still more prone to die, as seen at lower MMC concentrations. Treating the cells with IR, on the other hand, led to the specific induction of DSBs, which might require the function of FAN1 in a different step, downstream of end resection.

FAN1 was shown to be important for the recovery of stalled replication forks (164,176). The induction of FAN1/ICL-dependent DSBs could therefore also be related to such a function, where the replication fork is stalled in front of an ICL. In that case, the ICL would not necessarily be unhooked on both sides and the potential of FAN1 to bypass ICL-containing DNA would not be needed. Consequently, treating cells with other replication fork blocking agents should also lead to a decrease in DSBs upon FAN1 loss, which could be tested.

It is interesting that MUS81 was also shown to induce DSBs after the induction of ICLs. It also has a role in the recovery of stalled replication forks by cleaving intermediate structures (24,177). MUS81-dependent breaks actually seem to come from a persistence of stalled forks in the absence of XPF-SNM1A-dependent ICL unhooking (58). Like FAN1, MUS81 depletion did not affect the ability of *Xenopus* egg extracts to repair ICLs in a replication-dependent way. It could therefore be that the two proteins act in a common pathway. Indeed, I could show by PFGE and survival assays that the concomitant loss of both proteins did not alter the effect seen if only one protein was depleted. This means that the two proteins are not redundant and might act in the same pathway that resolves ICL-stalled replication forks.

In general, the cell biological studies of unhooking functions are lagging behind the biochemistry. To take these studies further, one should look more closely at redundancies between different potential ICL unhooking enzymes and other nucleases involved in ICL repair to take these studies further. Furthermore, ICL disappearance, for example with digoxigenin-tagged psoralen, could be used to study ICL unhooking enzymes (130). Another way of showing the unhooking of ICLs in cells would potentially be by performing 2-D gel analysis containing a native and a denaturing dimension on sheared chromosomal DNA so that hooked and unhooked DNA would run differently. As mentioned, separation-of function mutants would be useful to test FAN1's role in unhooking, independently of its other roles.

3.3 Evidence for FAN1 involvement in processes other than ICL unhooking

In recent years it has become clear that FAN1 is a versatile protein, having functions in different DNA metabolic pathways. As it seems to be redundant with other nucleases in each of these functions, it has been difficult to study them.

The sensitivity of FAN1-depleted cells to ICL agents is not as pronounced as upon down-regulation of other factors, such as FANC proteins. For its function in other DNA repair processes, no indicative sensitivity is observed. FAN1 depletion does not lead to hypersensitivity to CPT, for example, which induces a large amount of DSBs. The situation is similar when using replication fork stalling agents, such HU or APH. However, in specific assays, like the DR-GFP reporter assay to assess DSB repair, FAN1 depletion shows a mild phenotype (158,159). It could be that these assays are more sensitive and can detect minor changes in the phenotype, which are not observable by survival studies. This would mean that FAN1 has only minor functions in these pathways, maybe acting as a 'back-up' nuclease, which comes into play when the main nucleases responsible for the task are missing. Moreover, depending on the design of the reporter assays, there could already be some channeling into a specific repair pathway, leading to a certain bias in the results.

3.3.1 Replication fork recovery

FAN1 was proposed to act independently of the UBZ domain in ICL repair and to have a second, UBZ-dependent function in the recovery of stalled replication forks (176). Intriguingly, MEFs expressing a FAN1 variant with a mutated UBZ domain were not sensitive to any particular drug, but showed an increase in chromosomal aberrations upon treatment with HU or MMC. The authors correlate the UBZ-defective mutant phenotypes with those resulting from FANCD2 loss and therefore conclude that this interaction is responsible for the function in prevention of chromosomal instability. However, it is unclear why no sensitivity is detected upon the treatment with such agents, given the high amount of chromosomal abnormalities in these cells. It could be that the effects of the UBZ FAN1 mutant and FANCD2 deficiency are independent and that FAN1 actually depends on the presence of another (ubiquitinated) protein for recruitment. It would still work in a common pathway with FANCD2, as the two were shown to be epistatic, but the UBZ domain would actually be required for the binding to another protein. In fact, the authors did not show the direct interaction of FAN1 and FANCD2 upon replication fork stalling. One interaction candidate could be PCNA, as this travels along with the replication fork and becomes ubiquitinated upon fork stalling. Furthermore, in a different publication, FAN1 was already shown to be required for the recovery of stalled replication forks upon APH treatment, but there this function was independent of the UBZ domain or the FA core complex (164).

It is in general very difficult to differentiate between the functions of replication fork recovery and ICL unhooking, as an ICL is a special type of replication fork stalling agent and unhooking is essentially a special form of replication fork recovery.

3.3.2 Homologous recombination

FAN1 depletion was shown to decrease the efficiency of HR, but so far no further study was undertaken to pinpoint FAN1 to a particular step in HR (158,159). I previously tested the possibility of FAN1 being a HJ resolvase, even though such an activity is low *in vitro*. It could

however be that this is enhanced *in vivo* by specific regulation mechanisms. For this, I co-depleted known HJ resolvases from HEK293 cells that contain the DR-GFP plasmid, allowing for HR scoring as a function of the fluorescence of the cells. None of the co-depleted proteins (GEN1, MUS81, XPF, SLX1 or MLH1) led to markedly lower HR efficiency compared to the single depletions, while RAD51 depletion, serving as a control, always showed a striking effect. Again, as is the case for replication fork stalling agents, FAN1 depletion does not sensitize cells to drugs or radiation that induce DSBs. It is highly probable therefore that another, redundant, protein fulfills a similar function, and the effect is only seen when using specific, more sensitive assays and not in overall cell survival. Due to the lack of additivity with other known HJ resolvases, it is unlikely that FAN1 has its main function in HJ resolution, although it cannot be excluded yet. Other potential substrates for FAN1 in the cell are D-loops, as these were cleaved very efficiently *in vitro,* an activity that has been attributed also to the MUS81-EME2 (but not MUS81-EME1) complex (158,178). D-loop cleavage should however rather have an anti-recombinational effect in the cell, as it inhibits strand invasion and therefore further HR steps. Intriguingly, the MUS81-EME2 complex is also important for the recovery of stalled replication forks and its loss leads to an increase in chromosomal abnormalities, but it is not involved in the resolution of HJs, which is mediated by MUS81-EME1 (179). Maybe we can learn from the studies that have been done with MUS81-EME2 and test whether FAN1's function in HR and/or replication fork recovery is similar to the one of MUS81-EME2 or whether the two proteins act in completely different ways. MUS81-FAN1 co-depletion should have given an increased effect in the DR-GFP reporter assay if they really act in a similar and partial redundant manner, but this could be tested again by specifically disrupting EME2. The two functions of FAN1 in stalled replication fork recovery and HR could actually be the same, as FAN1 might act in a pathway that recovers stalled replication forks in a recombination-dependent manner.

3.4 FAN1 and FEN1

We have tested the relative abundance and activities of FAN1 and FEN1 and looked for a potential link and redundancies between the two proteins. Although there is still a need for further experiments to draw final conclusions, it seems that the two proteins could act in a partially redundant fashion upon treating cells with MMC. FAN1 expression was slightly increased upon FEN1 loss both at the transcriptional and protein level, probably due to a higher load of specific repair intermediates. FEN1 has been shown to be able to resolve stalled replication forks involving its GEN activity and a specific interaction with WRN (152). Such an activity could also be used during ICL repair. Given that the ICL is more than 2 bps away from the junction, 5' flap cleavage could occur, as shown by my results. Such a cleaved structure could then further be used by FAN1 or SNM1A to perform the bypass reaction over the ICL.

FEN1 is an essential protein and its deficiency is lethal in human cells and leads to a slow growth rate in DT40 chicken cells. Therefore, the possibility that the observed MMC sensiticity

is due to the general sickness of the cells rather than to the presence of the ICLs cannot be excluded. For example, the decreased growth rate, which includes slower replication, could hinder the replication-dependent ICL repair pathway, which could in turn lead to the higher sensitivity of the cells to this type of damage.

3.5 FAN1 and SNM1A

SNM1A is crucial during replication-dependent and -independent ICL repair, where its function was proposed to lie in the exonucleolytic degradation of the oligonucleotide around a cross-link.

Interestingly, SNM1A seems to stop a short distance from the incision 3' to the ICL, possibly due to a steric inhibition of further exonucleolytic degradation by the remnant of the ICL (58,180). In contrast, the exonucleolytic activity of FAN1 is unaffected by the ICL. This could be explained by its +3 scissile phosphate register, which makes it possible to bypass the lesion (58). The recently identified crystal structure of SNM1A catalytic domain reveals a putative DNA-binding groove (180). Mutation of two lysine residues at one end of this groove (K904 and K906) led to a decreased processivity of the enzyme. Interestingly, the ability of SNM1A to digest past an ICL was affected to a greater extent than the activity on native DNA substrates. This points towards the importance of this groove to accommodate damaged and distorted DNA structures.

SNM1A and FAN1 have both the ability to digest ICL-containing DNA, which seems to be a requirement for further lesion bypass by TLS pols. It could therefore be assumed that they have (partially) redundant functions in the cell. To test this, one could perform double depletion of the two proteins, where a stronger phenotype would be expected than upon the loss of the single proteins. Furthermore, it would be interesting to see under what circumstances one or the other protein comes into play and whether they have different propensities to process S-phase versus non S-phase dependent structures. As SNM1A has no endonucleolytic activity, it could also be that FAN1, or yet another nuclease, generates the first cut as an entry point for SNM1A, a function that has been proposed for XPF (58).

3.6 Potential parallel pathways to repair ICLs

It is likely that at least two different pathways to repair ICLs exist, a FA-dependent and a FA-independent one. The question of course is, which pathway comes into play when and under what conditions. FAN1 seems to have a function in ICL repair that is independent of the FA pathway. FAN1 UBZ domain was dispensable for its function in ICL repair and thus the possibility of an alternative recruitment to sites of ICL damage exists (176). Also the fact that prokaryotic and lower eukaryotic FAN1 do not possess a UBZ domain means that this domain has evolved later for a more specific recruitment and higher specificity, but that certain functions can already be achieved by FAN1 independently of the UBZ domain. If FAN1 function is independent of the FA core complex during the repair of ICLs, even though it

seems to be recruited in a monoubiquitinated FANCD2 (FANCD2-Ub)-dependent manner, the question arises why this specific recruitment happens at all. It could be that FAN1 could sense the damage and be recruited to it already by itself or by the interaction with a different (ubiquitinated or not) protein and that FANCD2-Ub somehow reinforces this interaction, but is not essential for its function there. Furthermore, it could be that this interaction is only required under certain conditions, for example in a specific cell cycle phase. To test this, the function of FAN1 outside of S-phase would have to be assessed. It is also possible that ubiquitin-independent interactions between FAN1 and FANCD2 exist. In line with that, Kratz et al. showed that the wild-type version of FAN1 was able to associate with both, unmodified FANCD2 and FANCD2-Ub, while the UBZ mutant only pulled-down the unmodified version. The results from the other primary papers do not all fully agree with this notion (157,159,160).

Mutations in genes of the FA pathway lead to the severe genetic disorder Fanconi anemia, which predisposes affected individuals to cancers, like lymphoma and squamous cell carcinoma. FAN1 deficiency has, on the other hand, been associated with pancreatic and colon cancer and the kidney disease KIN, and its involvement in a FA-independent pathway could explain these discrepancies in symptoms (171,173,174). This means that one or the other pathway could be preferably used in certain tissues, cell cycle phases, or to repair different types of ICL damage based on their structures (171). The nuclease activity of different unhooking nucleases would have to be tested on a variety of ICLs to draw further conclusions about this point. Based on available evidence, XPF would be the preferred candidate for a FA-dependent nuclease (56).

There are several possibilities how an ICL could be unhooked by the known proteins. XPF could cleave 5' and 3' to the ICL if the replication fork is close to the ICL. Further processing could be performed by SNM1A or FAN1. SLX1, FEN1 or FAN1 could also provide an entry point 5' from the lesion at a stalled replication fork and MUS81 might cleave 3' to it, while FAN1 and SNM1A again could perform the degradation over the cross-link. In a replication-independent situation, XPF could provide the entry point again for FAN1- or SNM1A-dependent degradation.

4 Conclusions and perspectives

DNA repair mechanisms are essential to prevent genome instability and cancer. Despite the very fast accumulation of knowledge in this exiting field over the past decades, many details still remain unclear. The rapid growth of -omics data generated in high-throughput experiments permits the study of a variety of processes in the cell at once. Nevertheless, the thorough mechanistic understanding of the given processes is often fragmentary or lacking completely. In an attempt to learn more about FAN1-dependent processes in the cell, specifically about its role in ICL repair, I studied its ability to bypass such lesions *in vitro*. Even though it had become clear, also thanks to other studies, that FAN1 is able to perform such a bypass, we were still not able to unambiguously show that this function is also employed by the cell to repair this kind of damage. To show this, the easiest method would be to look at the disappearance of fluorescently tagged versions of ICLs in the presence or absence of FAN1. It would have to be designed in a way that the unhooked ICL no longer fluoresces and thus the actual release from one strand of the DNA could already be assessed. Other remaining questions are whether FAN1 only acts during S-phase or whether it could also have a G1-specific function, and, of course, how this is regulated in the cell and which pathways are involved there. Specific ICL induction during G1 phase has already been achieved, using a psoralen analogue that can be fluorescently tagged by an antibody, and this method could be useful for the study of such a function (130). Of course one major question concerning FAN1 still remains: the importance of the UBZ domain and whether this is used for the interaction with proteins other than FANCD2-Ub. Furthermore, it is still unclear if and how FAN1 can function in such diverse DNA metabolic pathways such as ICL repair, HR, stalled replication fork recovery, or MMR. One useful tool to study each of these functions would be separation-of-function mutations or specific interaction site mutants that abolish the association with specific partners from each pathway. Such interaction partners could potentially be identified by pull-downs and mass spectrometry after treating cells with different types of drugs. Other nucleases can most likely complement at least some of FAN1's roles in each of these pathways. The different pathways are furthermore often related to each other and are not so easy to separate completely, which will make this study complicated and will undoubtedly keep many researchers from different laboratories busy for some time. Nevertheless, studying FAN1 in particular and DNA repair processes in general is an important task, as further knowledge in the field could potentially lead to improved cancer chemotherapy and decreased side-effects in patients.

5 Materials and Methods to section 2.2

5.1 Substrates for endonuclease assays

Substrate production was carried out as indicated in (181), using the oligonucleotides therein.

5.2 Culturing of DT40 chicken cells

Chicken B lymphocyte DT40 cells were cultured in RPMI-1640 medium (GIBCO), supplemented with 10% heat-inactivated Fetal Bovine Serum (GIBCO), 1% chicken serum (GIBCO), 10 µM β-Mercaptoethanol (GIBCO), 10000 U/ml penicillin and 10 mg/ml streptomycin (GIBCO) and kept at 39.5°C with 5% CO_2.

5.3 Culturing of TSCER2 cell lines and nuclear extract preparation

Cells were kept in RPMI-1640 medium (GIBCO), supplemented with 5% heat-inactivated Fetal Bovine Serum (GIBCO), 18 mM Sodium Pyruvate (GIBCO), 10000 U/ml penicillin and 10 mg/ml streptomycin (GIBCO) and at 37°C in 5% CO_2. The Cas9-mediated knock-out cell lines were obtained and kindly provided by Dr. Shunsuke Kobayashi. Nuclear extracts of TSCER2 or HEK293 cells were made as described previously in (182).

5.4 FAN1 overexpression in DT40 FEN1 KO cells

The DT40 FEN1 KO cell lines were made by Dr. Saho Kobayashi as described in (156). FAN1 chicken cDNA was then cloned into the GFP-expressing vector pIRES2-EGFP also as described in (156). The human FAN1 cDNA was directly tagged with GFP, by using the pEGFP-C1 plasmid as described in (158). The expression plasmid was linearized with PvuI prior to transfection into DT40 cells. 30 µg of linear targeting vector were used for electroporating 5 x 10^6 cells using a Gene Pulser (BioRad) at 550 V, 25 µF. The cells were then transferred to a culture dish in 20 ml medium and incubated for 16 h at 39.5°C. Subsequently, selection medium was added containing 2 mg/ml Geneticin (G418, GIBCO) and the cells were plated on 96-well plates (200 µl/well). After around 7 days colonies started to appear and were transferred to 24-well plates. After further expansion the cells were analyzed by FACS (CyAn ADP9 Cytometer, Beckman Coulter) for the expression of GFP and stored for future experiments.

5.5 Preparation of whole cell extracts from DT40 cells

2 x 10^8 DT40 cells were grown in 200 ml of medium overnight, harvested, washed with cold PBS and resuspended in isotonic lysis [10 mM Tris-HCl, pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 0.32 M sucrose, 1/100 EDTA-free protease inhibitor cocktail (Roche), 1 mM Dithiothreitol (DTT), 0.1 mM Spermine, 0.5 mM Spermidine]. IGEPAL CA-630 (Sigma) was then added to a

final concentration of 0.3%, followed by immediate centrifugation at full speed. The supernatant was aliquoted for activity assays.

5.6 Nuclease assays with DT40 extracts

50 µg of DT40 extract were incubated with 50 fmol of radiolabeled DNA substrate in a volume of 25 µl and 1 x nuclease buffer [25 mM Hepes-KOH, pH 7.4, 1 mM $MgCl_2$, 2.5% Polyethylenglycol (PEG), 0.5 mg/ml BSA, 30 mM DTT] for 30 min at room temperature. The reaction was stopped with 0.1% SDS, 14 mM EDTA, and 0.1 mg/ml Proteinase K for 15 min at 55°C. Loading buffer (80% formamide, 50 mM Tris-HCl pH 8, 1 mM EDTA) was added and the samples were separated on a 20% denaturing polyacrylamide gel containing 8 M Urea for 1 hr at 40 V/cm. The gel was dried and the bands were visualized on a PhosphoImager (Typhoon FLA 9500, GE healthcare).

5.7 Methylcellulose survival assay with DT40 cells

DT40 medium, as specified above, was supplemented with 1.5% methylcellulose (Sigma) and mixed under constant stirring at 4°C. MMC was added one day prior to the assay at the indicated concentrations and the MMC-containing methylcellulose was stirred at 4°C again overnight. 100 to 3000 cells/ml were mixed with the methylcellulose-containing medium in 6-well plates (35 mm diameter) depending on the expected survival and grown for one week before counting the colonies.

5.8 quantitative real-time (RT) PCR

qRT PCR was performed as described in (183). Primers used were the following:

GAPDH forward:	AGG GCT GCT TTT AAC TCT GGT
GAPDH reverse:	CCC CAC TTG ATT TTG GAG GGA
FAN1 forward:	TGG CCC CAG GAA GAA GAA ATT
FAN1 reverse:	GAT GTT CTG AGC AAT AAA ACA CAG GA
FEN1 forward:	CTG TGG ACC TCA TCC AGA AGC A
FEN1 reverse:	CCA GCA CCT CAG GTT CCA AGA
SNM1A forward:	AAT CAC TGT CCA GGT GCT GTC AT
SNM1A reverse:	TGG AAA GGT GTA TTC TGG GCT AC

5.9 Antibodies

Rabbit α-FEN1 (Bethyl laboratories, 1: 5000), mouse α-GAPDH (Millipore, 1:40'000), the remaining antibodies were as described in (181).

5.10 siRNAs

siFEN1: UAAGUCCAUUGUUACAUGAAA. Other siRNAs and treatments as described in (181).

6 References

1. Hoeijmakers, J. H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374
2. Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer. *Cell* **100**, 57-70
3. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell* **144**, 646-674
4. Reha-Krantz, L. J. (2010) DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim Biophys Acta* **1804**, 1049-1063
5. Jiricny, J. (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* **7**, 335-346
6. Yi, C., and He, C. (2013) DNA repair by reversal of DNA damage. *Cold Spring Harb Perspect Biol* **5**, a012575
7. Zharkov, D. O. (2008) Base excision DNA repair. *Cell Mol Life Sci* **65**, 1544-1565
8. Marteijn, J. A., Lans, H., Vermeulen, W., and Hoeijmakers, J. H. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* **15**, 465-481
9. Davis, A. J., and Chen, D. J. (2013) DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* **2**, 130-143
10. Jasin, M., and Rothstein, R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol* **5**, a012740
11. Krejci, L., Altmannova, V., Spirek, M., and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic Acids Res* **40**, 5795-5818
12. Clauson, C., Scharer, O. D., and Niedernhofer, L. (2013) Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb Perspect Biol* **5**, a012732
13. Chang, D. J., and Cimprich, K. A. (2009) DNA damage tolerance: when it's OK to make mistakes. *Nat Chem Biol* **5**, 82-90
14. Goodman, M. F., and Woodgate, R. (2013) Translesion DNA polymerases. *Cold Spring Harb Perspect Biol* **5**, a010363
15. Leman, A. R., and Noguchi, E. (2013) The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication. *Genes (Basel)* **4**, 1-32
16. Levikova, M., and Cejka, P. (2015) The Saccharomyces cerevisiae Dna2 can function as a sole nuclease in the processing of Okazaki fragments in DNA replication. *Nucleic Acids Res*
17. Mechali, M. (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol* **11**, 728-738
18. Garg, P., and Burgers, P. M. (2005) DNA polymerases that propagate the eukaryotic DNA replication fork. *Crit Rev Biochem Mol Biol* **40**, 115-128
19. Kunkel, T. A. (1988) Exonucleolytic proofreading. *Cell* **53**, 837-840
20. Zeman, M. K., and Cimprich, K. A. (2014) Causes and consequences of replication stress. *Nat Cell Biol* **16**, 2-9
21. Lambert, S., and Carr, A. M. (2013) Impediments to replication fork movement: stabilisation, reactivation and genome instability. *Chromosoma* **122**, 33-45

22. Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., and Burgers, P. M. (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* **47**, 980-986

23. Ying, S., Minocherhomji, S., Chan, K. L., Palmai-Pallag, T., Chu, W. K., Wass, T., Mankouri, H. W., Liu, Y., and Hickson, I. D. (2013) MUS81 promotes common fragile site expression. *Nat Cell Biol* **15**, 1001-1007

24. Naim, V., Wilhelm, T., Debatisse, M., and Rosselli, F. (2013) ERCC1 and MUS81-EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis. *Nat Cell Biol* **15**, 1008-1015

25. Bester, A. C., Roniger, M., Oren, Y. S., Im, M. M., Sarni, D., Chaoat, M., Bensimon, A., Zamir, G., Shewach, D. S., and Kerem, B. (2011) Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* **145**, 435-446

26. Shiloh, Y. (2003) ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer* **3**, 155-168

27. Yeeles, J. T., Poli, J., Marians, K. J., and Pasero, P. (2013) Rescuing stalled or damaged replication forks. *Cold Spring Harb Perspect Biol* **5**, a012815

28. Lopes, M., Foiani, M., and Sogo, J. M. (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol Cell* **21**, 15-27

29. Yeeles, J. T., and Marians, K. J. (2011) The Escherichia coli replisome is inherently DNA damage tolerant. *Science* **334**, 235-238

30. Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K. J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., and Lopes, M. (2012) Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol* **19**, 417-423

31. Betous, R., Mason, A. C., Rambo, R. P., Bansbach, C. E., Badu-Nkansah, A., Sirbu, B. M., Eichman, B. F., and Cortez, D. (2012) SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes Dev* **26**, 151-162

32. Thangavel, S., Berti, M., Levikova, M., Pinto, C., Gomathinayagam, S., Vujanovic, M., Zellweger, R., Moore, H., Lee, E. H., Hendrickson, E. A., Cejka, P., Stewart, S., Lopes, M., and Vindigni, A. (2015) DNA2 drives processing and restart of reversed replication forks in human cells. *J Cell Biol* **208**, 545-562

33. Dronkert, M. L., and Kanaar, R. (2001) Repair of DNA interstrand cross-links. *Mutat Res* **486**, 217-247

34. Huang, Y., and Li, L. (2013) DNA crosslinking damage and cancer - a tale of friend and foe. *Transl Cancer Res* **2**, 144-154

35. Scharer, O. D. (2005) DNA interstrand crosslinks: natural and drug-induced DNA adducts that induce unique cellular responses. *Chembiochem* **6**, 27-32

36. Angelov, T., Guainazzi, A., and Scharer, O. D. (2009) Generation of DNA interstrand cross-links by post-synthetic reductive amination. *Org Lett* **11**, 661-664

37. Auerbach, A. D. (2009) Fanconi anemia and its diagnosis. *Mutat Res* **668**, 4-10

38. Kottemann, M. C., and Smogorzewska, A. (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* **493**, 356-363

39. Wang, A. T., and Smogorzewska, A. (2015) SnapShot: Fanconi anemia and associated proteins. *Cell* **160**, 354-354 e351

40. Rickman, K. A., Lach, F. P., Abhyankar, A., Donovan, F. X., Sanborn, E. M., Kennedy, J. A., Sougnez, C., Gabriel, S. B., Elemento, O., Chandrasekharappa, S. C., Schindler, D., Auerbach, A. D., and Smogorzewska, A. (2015) Deficiency of UBE2T, the E2 Ubiquitin Ligase Necessary for FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia. *Cell Rep* **12**, 35-41

41. Long, D. T., Joukov, V., Budzowska, M., and Walter, J. C. (2014) BRCA1 promotes unloading of the CMG helicase from a stalled DNA replication fork. *Mol Cell* **56**, 174-185

42. Meetei, A. R., Medhurst, A. L., Ling, C., Xue, Y., Singh, T. R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C. G., Hoatlin, M., Joenje, H., de Winter, J. P., and Wang, W. (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* **37**, 958-963

43. Gari, K., Decaillet, C., Stasiak, A. Z., Stasiak, A., and Constantinou, A. (2008) The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Mol Cell* **29**, 141-148

44. Singh, T. R., Saro, D., Ali, A. M., Zheng, X. F., Du, C. H., Killen, M. W., Sachpatzidis, A., Wahengbam, K., Pierce, A. J., Xiong, Y., Sung, P., and Meetei, A. R. (2010) MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell* **37**, 879-886

45. Ciccia, A., Ling, C., Coulthard, R., Yan, Z., Xue, Y., Meetei, A. R., Laghmani el, H., Joenje, H., McDonald, N., de Winter, J. P., Wang, W., and West, S. C. (2007) Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell* **25**, 331-343

46. Deans, A. J., and West, S. C. (2011) DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**, 467-480

47. Kim, H., and D'Andrea, A. D. (2012) Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev* **26**, 1393-1408

48. Wang, X., Andreassen, P. R., and D'Andrea, A. D. (2004) Functional interaction of monoubiquitinated FANCD2 and BRCA2/FANCD1 in chromatin. *Mol Cell Biol* **24**, 5850-5862

49. Yamamoto, K. N., Kobayashi, S., Tsuda, M., Kurumizaka, H., Takata, M., Kono, K., Jiricny, J., Takeda, S., and Hirota, K. (2011) Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proc Natl Acad Sci U S A* **108**, 6492-6496

50. Kawasaki, K., Takahashi, M., Natori, M., and Shibata, T. (1991) DNA sequence recognition by a eukaryotic sequence-specific endonuclease, Endo.SceI, from Saccharomyces cerevisiae. *J Biol Chem* **266**, 5342-5347

51. Yang, W. (2011) Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys* **44**, 1-93

52. Yang, W., Lee, J. Y., and Nowotny, M. (2006) Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol Cell* **22**, 5-13

53. Ciccia, A., McDonald, N., and West, S. C. (2008) Structural and functional relationships of the XPF/MUS81 family of proteins. *Annu Rev Biochem* **77**, 259-287

54. De Silva, I. U., McHugh, P. J., Clingen, P. H., and Hartley, J. A. (2000) Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol Cell Biol* **20**, 7980-7990

55. Kuraoka, I., Kobertz, W. R., Ariza, R. R., Biggerstaff, M., Essigmann, J. M., and Wood, R. D. (2000) Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J Biol Chem* **275**, 26632-26636

56. Klein Douwel, D., Boonen, R. A., Long, D. T., Szypowska, A. A., Raschle, M., Walter, J. C., and Knipscheer, P. (2014) XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol Cell* **54**, 460-471

57. Niedernhofer, L. J., Odijk, H., Budzowska, M., van Drunen, E., Maas, A., Theil, A. F., de Wit, J., Jaspers, N. G., Beverloo, H. B., Hoeijmakers, J. H., and Kanaar, R. (2004) The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Mol Cell Biol* **24**, 5776-5787

58. Wang, A. T., Sengerova, B., Cattell, E., Inagawa, T., Hartley, J. M., Kiakos, K., Burgess-Brown, N. A., Swift, L. P., Enzlin, J. H., Schofield, C. J., Gileadi, O., Hartley, J. A., and McHugh, P. J. (2011) Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes Dev* **25**, 1859-1870

59. Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar, R. (2006) The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. *Embo J* **25**, 4921-4932

60. Zhang, J., and Walter, J. C. (2014) Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair (Amst)* **19**, 135-142

61. Castor, D., Nair, N., Declais, A. C., Lachaud, C., Toth, R., Macartney, T. J., Lilley, D. M., Arthur, J. S., and Rouse, J. (2013) Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. *Mol Cell* **52**, 221-233

62. Kim, Y., Lach, F. P., Desetty, R., Hanenberg, H., Auerbach, A. D., and Smogorzewska, A. (2011) Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet* **43**, 142-146

63. Cattell, E., Sengerova, B., and McHugh, P. J. (2010) The SNM1/Pso2 family of ICL repair nucleases: from yeast to man. *Environ Mol Mutagen* **51**, 635-645

64. Demuth, I., Digweed, M., and Concannon, P. (2004) Human SNM1B is required for normal cellular response to both DNA interstrand crosslink-inducing agents and ionizing radiation. *Oncogene* **23**, 8611-8618

65. Bae, J. B., Mukhopadhyay, S. S., Liu, L., Zhang, N., Tan, J., Akhter, S., Liu, X., Shen, X., Li, L., and Legerski, R. J. (2008) Snm1B/Apollo mediates replication fork collapse and S Phase checkpoint activation in response to DNA interstrand cross-links. *Oncogene* **27**, 5045-5056

66. Gu, J., Li, S., Zhang, X., Wang, L. C., Niewolik, D., Schwarz, K., Legerski, R. J., Zandi, E., and Lieber, M. R. (2010) DNA-PKcs regulates a single-stranded DNA endonuclease activity of Artemis. *DNA Repair (Amst)* **9**, 429-437

67. Sengerova, B., Allerston, C. K., Abu, M., Lee, S. Y., Hartley, J., Kiakos, K., Schofield, C. J., Hartley, J. A., Gileadi, O., and McHugh, P. J. (2012) Characterization of the human SNM1A and SNM1B/Apollo DNA repair exonucleases. *J Biol Chem* **287**, 26254-26267

68. Hazrati, A., Ramis-Castelltort, M., Sarkar, S., Barber, L. J., Schofield, C. J., Hartley, J. A., and McHugh, P. J. (2008) Human SNM1A suppresses the DNA repair defects of yeast pso2 mutants. *DNA Repair (Amst)* **7**, 230-238

69. Ishiai, M., Kimura, M., Namikoshi, K., Yamazoe, M., Yamamoto, K., Arakawa, H., Agematsu, K., Matsushita, N., Takeda, S., Buerstedde, J. M., and Takata, M. (2004) DNA cross-link repair protein SNM1A interacts with PIAS1 in nuclear focus formation. *Mol Cell Biol* **24**, 10733-10741

70. Yang, K., Moldovan, G. L., and D'Andrea, A. D. (2010) RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger. *J Biol Chem* **285**, 19085-19091

71. Allerston, C. K., Lee, S. Y., Newman, J. A., Schofield, C. J., McHugh, P. J., and Gileadi, O. (2015) The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities. *Nucleic Acids Res*

72. Lehmann, A. R. (2006) Translesion synthesis in mammalian cells. *Exp Cell Res* **312**, 2673-2676

73. Chen, Y. W., Cleaver, J. E., Hanaoka, F., Chang, C. F., and Chou, K. M. (2006) A novel role of DNA polymerase eta in modulating cellular sensitivity to chemotherapeutic agents. *Mol Cancer Res* **4**, 257-265

74. Albertella, M. R., Green, C. M., Lehmann, A. R., and O'Connor, M. J. (2005) A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res* **65**, 9799-9806

75. Wang, X., Peterson, C. A., Zheng, H., Nairn, R. S., Legerski, R. J., and Li, L. (2001) Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. *Mol Cell Biol* **21**, 713-720

76. Ho, T. V., Guainazzi, A., Derkunt, S. B., Enoiu, M., and Scharer, O. D. (2011) Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. *Nucleic Acids Res* **39**, 7455-7464

77. Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I., and Lehmann, A. R. (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol Cell* **37**, 714-727

78. Minko, I. G., Harbut, M. B., Kozekov, I. D., Kozekova, A., Jakobs, P. M., Olson, S. B., Moses, R. E., Harris, T. M., Rizzo, C. J., and Lloyd, R. S. (2008) Role for DNA polymerase kappa in the processing of N2-N2-guanine interstrand cross-links. *J Biol Chem* **283**, 17075-17082

79. Gan, G. N., Wittschieben, J. P., Wittschieben, B. O., and Wood, R. D. (2008) DNA polymerase zeta (pol zeta) in higher eukaryotes. *Cell Res* **18**, 174-183

80. Nojima, J., Kuratsune, H., Suehisa, E., Iwatani, Y., and Kanakura, Y. (2005) Acquired activated protein C resistance associated with IgG antibodies against beta2-glycoprotein I and prothrombin as a strong risk factor for venous thromboembolism. *Clin Chem* **51**, 545-552

81. Lee, Y. S., Gregory, M. T., and Yang, W. (2014) Human Pol zeta purified with accessory subunits is active in translesion DNA synthesis and complements Pol eta in cisplatin bypass. *Proc Natl Acad Sci U S A* **111**, 2954-2959

82. Kim, H., Yang, K., Dejsuphong, D., and D'Andrea, A. D. (2012) Regulation of Rev1 by the Fanconi anemia core complex. *Nat Struct Mol Biol* **19**, 164-170

83. Shen, X., Jun, S., O'Neal, L. E., Sonoda, E., Bemark, M., Sale, J. E., and Li, L. (2006) REV3 and REV1 play major roles in recombination-independent repair of DNA interstrand cross-links mediated by monoubiquitinated proliferating cell nuclear antigen (PCNA). *J Biol Chem* **281**, 13869-13872

84. Raschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J. D., Ellenberger, T. E., Scharer, O. D., and Walter, J. C. (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969-980

85. Moldovan, G. L., Madhavan, M. V., Mirchandani, K. D., McCaffrey, R. M., Vinciguerra, P., and D'Andrea, A. D. (2010) DNA polymerase POLN participates in cross-link repair and homologous recombination. *Mol Cell Biol* **30**, 1088-1096

86. Takata, K., Shimizu, T., Iwai, S., and Wood, R. D. (2006) Human DNA polymerase N (POLN) is a low fidelity enzyme capable of error-free bypass of 5S-thymine glycol. *J Biol Chem* **281**, 23445-23455

87. Yamanaka, K., Minko, I. G., Takata, K., Kolbanovskiy, A., Kozekov, I. D., Wood, R. D., Rizzo, C. J., and Lloyd, R. S. (2010) Novel enzymatic function of DNA polymerase nu in translesion DNA synthesis past major groove DNA-peptide and DNA-DNA cross-links. *Chem Res Toxicol* **23**, 689-695

88. Weterings, E., and Chen, D. J. (2008) The endless tale of non-homologous end-joining. *Cell Res* **18**, 114-124

89. Symington, L. S. (2002) Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66**, 630-670, table of contents

90. McEachern, M. J., and Haber, J. E. (2006) Break-induced replication and recombinational telomere elongation in yeast. *Annu Rev Biochem* **75**, 111-135

91. Zhang, J. (2013) The role of BRCA1 in homologous recombination repair in response to replication stress: significance in tumorigenesis and cancer therapy. *Cell Biosci* **3**, 11

92. Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S. P. (2007) Human CtIP promotes DNA end resection. *Nature* **450**, 509-514

93. Cannavo, E., and Cejka, P. (2014) Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* **514**, 122-125

94. Garcia, V., Phelps, S. E., Gray, S., and Neale, M. J. (2011) Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature* **479**, 241-244

95. Lamarche, B. J., Orazio, N. I., and Weitzman, M. D. (2010) The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett* **584**, 3682-3695

96. Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J. L., Wyman, C., Modrich, P., and Kowalczykowski, S. C. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* **25**, 350-362

97. Eid, W., Steger, M., El-Shemerly, M., Ferretti, L. P., Pena-Diaz, J., Konig, C., Valtorta, E., Sartori, A. A., and Ferrari, S. (2010) DNA end resection by CtIP and exonuclease 1 prevents genomic instability. *EMBO Rep* **11**, 962-968

98. Heyer, W. D., Ehmsen, K. T., and Liu, J. (2010) Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44**, 113-139

99. Schwartz, E. K., and Heyer, W. D. (2011) Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes. *Chromosoma* **120**, 109-127

100. Svendsen, J. M., and Harper, J. W. (2010) GEN1/Yen1 and the SLX4 complex: Solutions to the problem of Holliday junction resolution. *Genes Dev* **24**, 521-536

101. Huhn, D., Bolck, H. A., and Sartori, A. A. (2013) Targeting DNA double-strand break signalling and repair: recent advances in cancer therapy. *Swiss Med Wkly* **143**, w13837

102. Helleday, T., Lo, J., van Gent, D. C., and Engelward, B. P. (2007) DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amst)* **6**, 923-935

103. Llorente, B., Smith, C. E., and Symington, L. S. (2008) Break-induced replication: what is it and what is it for? *Cell Cycle* **7**, 859-864

104. Chapman, J. R., Taylor, M. R., and Boulton, S. J. (2012) Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47**, 497-510

105. Ferretti, L. P., Lafranchi, L., and Sartori, A. A. (2013) Controlling DNA-end resection: a new task for CDKs. *Front Genet* **4**, 99

106. Huertas, P., Cortes-Ledesma, F., Sartori, A. A., Aguilera, A., and Jackson, S. P. (2008) CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature* **455**, 689-692

107. Lafranchi, L., de Boer, H. R., de Vries, E. G., Ong, S. E., Sartori, A. A., and van Vugt, M. A. (2014) APC/C(Cdh1) controls CtIP stability during the cell cycle and in response to DNA damage. *Embo J* **33**, 2860-2879

108. Yun, M. H., and Hiom, K. (2009) CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* **459**, 460-463

109. Kowalczykowski, S. C. (2015) An Overview of the Molecular Mechanisms of Recombinational DNA Repair. *Cold Spring Harb Perspect Biol* **7**

110. Bunting, S. F., Callen, E., Wong, N., Chen, H. T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., Cao, L., Xu, X., Deng, C. X., Finkel, T., Nussenzweig, M., Stark, J. M., and Nussenzweig, A. (2010) 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243-254

111. Bouwman, P., Aly, A., Escandell, J. M., Pieterse, M., Bartkova, J., van der Gulden, H., Hiddingh, S., Thanasoula, M., Kulkarni, A., Yang, Q., Haffty, B. G., Tommiska, J., Blomqvist, C., Drapkin, R., Adams, D. J., Nevanlinna, H., Bartek, J., Tarsounas, M., Ganesan, S., and Jonkers, J. (2010) 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nat Struct Mol Biol* **17**, 688-695

112. Nakanishi, K., Yang, Y. G., Pierce, A. J., Taniguchi, T., Digweed, M., D'Andrea, A. D., Wang, Z. Q., and Jasin, M. (2005) Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. *Proc Natl Acad Sci U S A* **102**, 1110-1115

113. Niedzwiedz, W., Mosedale, G., Johnson, M., Ong, C. Y., Pace, P., and Patel, K. J. (2004) The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell* **15**, 607-620

114. Kee, Y., and D'Andrea, A. D. (2010) Expanded roles of the Fanconi anemia pathway in preserving genomic stability. *Genes Dev* **24**, 1680-1694

115. Unno, J., Itaya, A., Taoka, M., Sato, K., Tomida, J., Sakai, W., Sugasawa, K., Ishiai, M., Ikura, T., Isobe, T., Kurumizaka, H., and Takata, M. (2014) FANCD2 binds CtIP and regulates DNA-end resection during DNA interstrand crosslink repair. *Cell Rep* **7**, 1039-1047

116. Murina, O., von Aesch, C., Karakus, U., Ferretti, L. P., Bolck, H. A., Hanggi, K., and Sartori, A. A. (2014) FANCD2 and CtIP cooperate to repair DNA interstrand crosslinks. *Cell Rep* **7**, 1030-1038

117. Adamo, A., Collis, S. J., Adelman, C. A., Silva, N., Horejsi, Z., Ward, J. D., Martinez-Perez, E., Boulton, S. J., and La Volpe, A. (2010) Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol Cell* **39**, 25-35

118. Pace, P., Mosedale, G., Hodskinson, M. R., Rosado, I. V., Sivasubramaniam, M., and Patel, K. J. (2010) Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* **329**, 219-223

119. Friedberg, E. C. (2001) How nucleotide excision repair protects against cancer. *Nat Rev Cancer* **1**, 22-33

120. Nouspikel, T. (2009) DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. *Cell Mol Life Sci* **66**, 994-1009

121. Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J. P., Minguillon, J., Ramirez, M. J., Pujol, R., Casado, J. A., Banos, R., Rio, P., Knies, K., Zuniga, S., Benitez, J., Bueren, J. A., Jaspers, N. G., Scharer, O. D., de Winter, J. P., Schindler, D., and Surralles, J. (2013) Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* **92**, 800-806

122. Kashiyama, K., Nakazawa, Y., Pilz, D. T., Guo, C., Shimada, M., Sasaki, K., Fawcett, H., Wing, J. F., Lewin, S. O., Carr, L., Li, T. S., Yoshiura, K., Utani, A., Hirano, A., Yamashita, S., Greenblatt, D., Nardo, T., Stefanini, M., McGibbon, D., Sarkany, R., Fassihi, H., Takahashi, Y., Nagayama, Y., Mitsutake, N., Lehmann, A. R., and Ogi, T. (2013) Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *Am J Hum Genet* **92**, 807-819

123. de Laat, W. L., Jaspers, N. G., and Hoeijmakers, J. H. (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev* **13**, 768-785

124. Kamileri, I., Karakasilioti, I., and Garinis, G. A. (2012) Nucleotide excision repair: new tricks with old bricks. *Trends Genet* **28**, 566-573

125. Knipscheer, P., Raschle, M., Scharer, O. D., and Walter, J. C. (2012) Replication-coupled DNA interstrand cross-link repair in Xenopus egg extracts. *Methods Mol Biol* **920**, 221-243

126. Zhang, J., Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., and Walter, J. C. (2015) DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol* **22**, 242-247

127. Budzowska, M., Graham, T. G., Sobeck, A., Waga, S., and Walter, J. C. (2015) Regulation of the Rev1-pol zeta complex during bypass of a DNA interstrand cross-link. *Embo J* **34**, 1971-1985

128. Huang, J., Liu, S., Bellani, M. A., Thazhathveetil, A. K., Ling, C., de Winter, J. P., Wang, Y., Wang, W., and Seidman, M. M. (2013) The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol Cell* **52**, 434-446

129. Thazhathveetil, A. K., Liu, S. T., Indig, F. E., and Seidman, M. M. (2007) Psoralen conjugates for visualization of genomic interstrand cross-links localized by laser photoactivation. *Bioconjug Chem* **18**, 431-437

130. Muniandy, P. A., Thapa, D., Thazhathveetil, A. K., Liu, S. T., and Seidman, M. M. (2009) Repair of laser-localized DNA interstrand cross-links in G1 phase mammalian cells. *J Biol Chem* **284**, 27908-27917

131. Enoiu, M., Jiricny, J., and Scharer, O. D. (2012) Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res* **40**, 8953-8964

132. Furuta, T., Ueda, T., Aune, G., Sarasin, A., Kraemer, K. H., and Pommier, Y. (2002) Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. *Cancer Res* **62**, 4899-4902

133. Zheng, H., Wang, X., Warren, A. J., Legerski, R. J., Nairn, R. S., Hamilton, J. W., and Li, L. (2003) Nucleotide excision repair- and polymerase eta-mediated error-prone removal of mitomycin C interstrand cross-links. *Mol Cell Biol* **23**, 754-761

134. Chen, Z., Xu, X. S., Yang, J., and Wang, G. (2003) Defining the function of XPC protein in psoralen and cisplatin-mediated DNA repair and mutagenesis. *Carcinogenesis* **24**, 1111-1121

135. Zhao, J., Jain, A., Iyer, R. R., Modrich, P. L., and Vasquez, K. M. (2009) Mismatch repair and nucleotide excision repair proteins cooperate in the recognition of DNA interstrand crosslinks. *Nucleic Acids Res* **37**, 4420-4429

136. Li, L., Peterson, C. A., Lu, X., Wei, P., and Legerski, R. J. (1999) Interstrand cross-links induce DNA synthesis in damaged and undamaged plasmids in mammalian cell extracts. *Mol Cell Biol* **19**, 5619-5630

137. Li, L., Peterson, C. A., Zhang, X., and Legerski, R. J. (2000) Requirement for PCNA and RPA in interstrand crosslink-induced DNA synthesis. *Nucleic Acids Res* **28**, 1424-1427

138. Smeaton, M. B., Hlavin, E. M., McGregor Mason, T., Noronha, A. M., Wilds, C. J., and Miller, P. S. (2008) Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. *Biochemistry* **47**, 9920-9930

139. Kumaresan, K. R., Hwang, M., Thelen, M. P., and Lambert, M. W. (2002) Contribution of XPF functional domains to the 5' and 3' incisions produced at the site of a psoralen interstrand cross-link. *Biochemistry* **41**, 890-896

140. Sharma, S., and Canman, C. E. (2012) REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. *Environ Mol Mutagen* **53**, 725-740

141. Wang, S., Liu, K., Xiao, L., Yang, L., Li, H., Zhang, F., Lei, L., Li, S., Feng, X., Li, A., and He, J. (2015) Characterization of a novel DNA glycosylase from S. sahachiroi involved in the reduction and repair of azinomycin B induced DNA damage. *Nucleic Acids Res*

142. Mullins, E. A., Shi, R., Parsons, Z. D., Yuen, P. K., David, S. S., Igarashi, Y., and Eichman, B. F. (2015) The DNA glycosylase AlkD uses a non-base-flipping mechanism to excise bulky lesions. *Nature* **527**, 254-258

143. Ward, T. A., Dudasova, Z., Sarkar, S., Bhide, M. R., Vlasakova, D., Chovanec, M., and McHugh, P. J. (2012) Components of a Fanconi-like pathway control Pso2-independent DNA interstrand crosslink repair in yeast. *PLoS Genet* **8**, e1002884

144. McHugh, P. J., Ward, T. A., and Chovanec, M. (2012) A prototypical Fanconi anemia pathway in lower eukaryotes? *Cell Cycle* **11**, 3739-3744

145. Daee, D. L., and Myung, K. (2012) Fanconi-like crosslink repair in yeast. *Genome Integr* **3**, 7

146. Daee, D. L., Ferrari, E., Longerich, S., Zheng, X. F., Xue, X., Branzei, D., Sung, P., and Myung, K. (2012) Rad5-dependent DNA repair functions of the Saccharomyces cerevisiae FANCM protein homolog Mph1. *J Biol Chem* **287**, 26563-26575

147. Karanja, K. K., and Livingston, D. M. (2009) C-terminal flap endonuclease (rad27) mutations: lethal interactions with a DNA ligase I mutation (cdc9-p) and suppression by proliferating cell nuclear antigen (POL30) in Saccharomyces cerevisiae. *Genetics* **183**, 63-78

148. Brosh, R. M., Jr., von Kobbe, C., Sommers, J. A., Karmakar, P., Opresko, P. L., Piotrowski, J., Dianova, I., Dianov, G. L., and Bohr, V. A. (2001) Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity. *Embo J* **20**, 5791-5801

149. Sakurai, S., Kitano, K., Morioka, H., and Hakoshima, T. (2008) Crystallization and preliminary crystallographic analysis of the catalytic domain of human flap endonuclease 1 in complex with a nicked DNA product: use of a DPCS kit for efficient protein-DNA complex crystallization. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 39-43

150. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A., and Tainer, J. A. (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198-211

151. Zhou, C., Pourmal, S., and Pavletich, N. P. (2015) Dna2 nuclease-helicase structure, mechanism and regulation by Rpa. *Elife* **4**

152. Zheng, L., Zhou, M., Chai, Q., Parrish, J., Xue, D., Patrick, S. M., Turchi, J. J., Yannone, S. M., Chen, D., and Shen, B. (2005) Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. *EMBO Rep* **6**, 83-89

153. Shen, B., Singh, P., Liu, R., Qiu, J., Zheng, L., Finger, L. D., and Alas, S. (2005) Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases. *Bioessays* **27**, 717-729

154. Zheng, L., Dai, H., Zhou, M., Li, M., Singh, P., Qiu, J., Tsark, W., Huang, Q., Kernstine, K., Zhang, X., Lin, D., and Shen, B. (2007) Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat Med* **13**, 812-819

155. Cannavo, E., Gerrits, B., Marra, G., Schlapbach, R., and Jiricny, J. (2007) Characterization of the interactome of the human MutL homologues MLH1, PMS1, and PMS2. *J Biol Chem* **282**, 2976-2986

156. Yoshikiyo, K., Kratz, K., Hirota, K., Nishihara, K., Takata, M., Kurumizaka, H., Horimoto, S., Takeda, S., and Jiricny, J. (2010) KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. *Proc Natl Acad Sci U S A* **107**, 21553-21557

157. Smogorzewska, A., Desetty, R., Saito, T. T., Schlabach, M., Lach, F. P., Sowa, M. E., Clark, A. B., Kunkel, T. A., Harper, J. W., Colaiacovo, M. P., and Elledge, S. J. (2010)

A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell* **39**, 36-47

158. Kratz, K., Schopf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavo, E., Sartori, A. A., Hengartner, M. O., and Jiricny, J. (2010) Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell* **142**, 77-88

159. MacKay, C., Declais, A. C., Lundin, C., Agostinho, A., Deans, A. J., MacArtney, T. J., Hofmann, K., Gartner, A., West, S. C., Helleday, T., Lilley, D. M., and Rouse, J. (2010) Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell* **142**, 65-76

160. Liu, T., Ghosal, G., Yuan, J., Chen, J., and Huang, J. (2010) FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science* **329**, 693-696

161. Desai, A., and Gerson, S. (2014) Exo1 independent DNA mismatch repair involves multiple compensatory nucleases. *DNA Repair (Amst)* **21**, 55-64

162. Lai, F., Hu, K., Wu, Y., Tang, J., Sang, Y., Cao, J., and Kang, T. (2012) Human KIAA1018/FAN1 nuclease is a new mitotic substrate of APC/C(Cdh1). *Chin J Cancer* **31**, 440-448

163. Shereda, R. D., Machida, Y., and Machida, Y. J. (2010) Human KIAA1018/FAN1 localizes to stalled replication forks via its ubiquitin-binding domain. *Cell Cycle* **9**, 3977-3983

164. Chaudhury, I., Stroik, D. R., and Sobeck, A. (2014) FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Mol Cell Biol* **34**, 3939-3954

165. Fontebasso, Y., Etheridge, T. J., Oliver, A. W., Murray, J. M., and Carr, A. M. (2013) The conserved Fanconi anemia nuclease Fan1 and the SUMO E3 ligase Pli1 act in two novel Pso2-independent pathways of DNA interstrand crosslink repair in yeast. *DNA Repair (Amst)* **12**, 1011-1023

166. Gwon, G. H., Kim, Y., Liu, Y., Watson, A. T., Jo, A., Etheridge, T. J., Yuan, F., Zhang, Y., Carr, A. M., and Cho, Y. (2014) Crystal structure of a Fanconi anemia-associated nuclease homolog bound to 5' flap DNA: basis of interstrand cross-link repair by FAN1. *Genes Dev* **28**, 2276-2290

167. Pennell, S., Declais, A. C., Li, J., Haire, L. F., Berg, W., Saldanha, J. W., Taylor, I. A., Rouse, J., Lilley, D. M., and Smerdon, S. J. (2014) FAN1 activity on asymmetric repair intermediates is mediated by an atypical monomeric virus-type replication-repair nuclease domain. *Cell Rep* **8**, 84-93

168. Wang, R., Persky, N. S., Yoo, B., Ouerfelli, O., Smogorzewska, A., Elledge, S. J., and Pavletich, N. P. (2014) DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science* **346**, 1127-1130

169. Zhao, Q., Xue, X., Longerich, S., Sung, P., and Xiong, Y. (2014) Structural insights into 5' flap DNA unwinding and incision by the human FAN1 dimer. *Nat Commun* **5**, 5726

170. Trujillo, J. P., Mina, L. B., Pujol, R., Bogliolo, M., Andrieux, J., Holder, M., Schuster, B., Schindler, D., and Surralles, J. (2012) On the role of FAN1 in Fanconi anemia. *Blood* **120**, 86-89

171. Zhou, W., Otto, E. A., Cluckey, A., Airik, R., Hurd, T. W., Chaki, M., Diaz, K., Lach, F. P., Bennett, G. R., Gee, H. Y., Ghosh, A. K., Natarajan, S., Thongthip, S., Veturi, U., Allen, S. J., Janssen, S., Ramaswami, G., Dixon, J., Burkhalter, F., Spoendlin, M., Moch, H., Mihatsch, M. J., Verine, J., Reade, R., Soliman, H., Godin, M., Kiss, D., Monga, G., Mazzucco, G., Amann, K., Artunc, F., Newland, R. C., Wiech, T., Zschiedrich, S., Huber, T. B., Friedl, A., Slaats, G. G., Joles, J. A., Goldschmeding, R., Washburn, J., Giles, R. H., Levy, S., Smogorzewska, A., and Hildebrandt, F. (2012) FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nat Genet* **44**, 910-915

172. Ionita-Laza, I., Xu, B., Makarov, V., Buxbaum, J. D., Roos, J. L., Gogos, J. A., and Karayiorgou, M. (2014) Scan statistic-based analysis of exome sequencing data identifies FAN1 at 15q13.3 as a susceptibility gene for schizophrenia and autism. *Proc Natl Acad Sci U S A* **111**, 343-348

173. Segui, N., Mina, L. B., Lazaro, C., Sanz-Pamplona, R., Pons, T., Navarro, M., Bellido, F., Lopez-Doriga, A., Valdes-Mas, R., Pineda, M., Guino, E., Vidal, A., Soto, J. L., Caldes, T., Duran, M., Urioste, M., Rueda, D., Brunet, J., Balbin, M., Blay, P., Iglesias, S., Garre, P., Lastra, E., Sanchez-Heras, A. B., Valencia, A., Moreno, V., Pujana, M. A., Villanueva, A., Blanco, I., Capella, G., Surralles, J., Puente, X. S., and Valle, L. (2015) Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. *Gastroenterology*

174. Smith, A. L., Alirezaie, N., Connor, A., Chan-Seng-Yue, M., Grant, R., Selander, I., Bascunana, C., Borgida, A., Hall, A., Whelan, T., Holter, S., McPherson, T., Cleary, S., Petersen, G. M., Omeroglu, A., Saloustros, E., McPherson, J., Stein, L. D., Foulkes, W. D., Majewski, J., Gallinger, S., and Zogopoulos, G. (2015) Candidate DNA repair susceptibility genes identified by exome sequencing in high-risk pancreatic cancer. *Cancer Lett*

175. Orans, J., McSweeney, E. A., Iyer, R. R., Hast, M. A., Hellinga, H. W., Modrich, P., and Beese, L. S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145**, 212-223

176. Lachaud, C., Moreno, A., Marchesi, F., Toth, R., Blow, J. J., and Rouse, J. (2016) Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Science*

177. Roseaulin, L., Yamada, Y., Tsutsui, Y., Russell, P., Iwasaki, H., and Arcangioli, B. (2008) Mus81 is essential for sister chromatid recombination at broken replication forks. *Embo J* **27**, 1378-1387

178. Pepe, A., and West, S. C. (2014) Substrate specificity of the MUS81-EME2 structure selective endonuclease. *Nucleic Acids Res* **42**, 3833-3845

179. Pepe, A., and West, S. C. (2014) MUS81-EME2 promotes replication fork restart. *Cell Rep* **7**, 1048-1055

180. Allerston, C. K., Lee, S. Y., Newman, J. A., Schofield, C. J., McHugh, P. J., and Gileadi, O. (2015) The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities. *Nucleic Acids Res* **43**, 11047-11060

181. Pizzolato, J., Mukherjee, S., Scharer, O. D., and Jiricny, J. (2015) FANCD2-associated Nuclease 1, but Not Exonuclease 1 or Flap Endonuclease 1, Is Able to Unhook DNA Interstrand Cross-links in Vitro. *J Biol Chem* **290**, 22602-22611

182. Repmann, S., Olivera-Harris, M., and Jiricny, J. (2015) Influence of oxidized purine processing on strand directionality of mismatch repair. *J Biol Chem* **290**, 9986-9999

183. Schmid, C. A., Robinson, M. D., Scheifinger, N. A., Muller, S., Cogliatti, S., Tzankov, A., and Muller, A. (2015) DUSP4 deficiency caused by promoter hypermethylation drives JNK signaling and tumor cell survival in diffuse large B cell lymphoma. *J Exp Med* **212**, 775-792

7 Acknowledgements

First of all, I wish to thank Prof. Joe Jiricny for his continuous support during my PhD. Even after so many years in research, he has not lost his curiosity and passion for molecular biology and such complex pathways as DNA repair. He looks out for his people and has made the IMCR such an amazing place to work at as it is now. I am grateful to have been a part of this lovely and inspiring community and to be one of the last members of Joe's lab and the institute under his supervision.

Next, a big thank you goes to my committee members, Prof. Orlando Schärer, Prof. Alex Sartori, and Prof. Lumir Krejci for the useful comments and discussions during the committee meetings. A special thank you goes to Orlando for the fruitful collaboration and for providing the cross-linked substrates, which were an essential part of my studies. I thank Alex for believing in me since my early days at the IMCR. Thank you also to Orlando and Lumir for making the long way to Zurich for my committee meetings and PhD defence.

A huge thank you goes to my second family, the Jiricny lab. I feel so blessed having such pleasant and caring work colleagues. Even if the project was not going well, I always enjoyed coming to the lab, because I knew you would be there and cheer me up. Special and heartfelt thanks go to Svenja, for her kindness and patience, introducing me to all the techniques in the lab and continuous moral support. Thanks to Tine, Anja, Lia, Martin, Mariela, Saho, Rachel, Simone and Svenja for the joyful lunch and coffee breaks, to the Joe's angels (including Martin, Cosimo and Hannah) for fun moments during our free time. Many thanks also to Anna, for her short but memorable time in our lab and a lot of fun moments. Another special thank goes to Antonio, for being such an enjoyable neighbor in 'the bay' and for always being open for scientific and non-scientific discussions.

Furthermore, I thank all the members of the IMCR for creating such an inspiring and pleasant working atmosphere and for always being very helpful when I was in need of help, advice or reagents. A special thanks goes to the angels of the institute, Farah, Odete, Paula and all the technicians for such a great job at making our everyday lab work so much easier.

Last but not least, I would like to express my biggest thanks to my family, who always support me and believe in me. To my grandparents, for always being curious and interested in my work, to my parents, sisters and brother for being the craziest and most amazing family I could have wished for, and to Julien for caring, supporting and put up with me during the whole time of my PhD.

8 Curriculum Vitae